(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
2 September 2004 (02.09.2004)

PCT

(10) International Publication Number
WO 2004/073657 A2

(51) International Patent Classification[7]: A61K

(21) International Application Number:
PCT/US2004/005455

(22) International Filing Date: 19 February 2004 (19.02.2004)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/448,784 19 February 2003 (19.02.2003) US

(71) Applicant *(for all designated States except US)*: **PROTEIN DESIGN LABS, INC.** [US/US]; 34801 Campus Drive, Fremont, CA 94555 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **AZIZ, Natasha** [US/US]; 411 California Avenue #5, Palo Alto, CA 94306 (US). **GISH, Kurt, C.** [US/US]; 37 Artuna Avenue, Piedmont, CA 94611 (US). **WILSON, Keith, E.** [TZ/US]; 219 Jeter Street, Redwood City, CA 94062 (US). **ZLOTNIK, Albert** [US/US]; 507 Alger Drive, Palo Alto, CA 94306 (US).

(74) Agents: **HALLUIN, Albert** et al.; 301 Ravenswood Avenue, Menlo Park, CA 94025 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: METHODS OF DIAGNOSIS OF CANCER, COMPOSITION AND METHODS OF SCREENING FOR MODULATORS OF CANCER

(57) Abstract: Described herein are genes whose expression are up-regulated or down-regulated in specific cancers or other diseases, or are otherwise regulated in disease. Related methods and compositions that can be used for diagnosis, prognosis, and treatment of those medical conditions are disclosed. Also described herein are methods that can be used to identify modulators of these selected conditions.

WO 2004/073657 A2

METHODS OF DIAGNOSIS OF CANCER, COMPOSITIONS AND METHODS OF SCREENING FOR MODULATORS OF CANCER

FIELD OF THE INVENTION

The invention relates to the identification of nucleic acid and protein expression profiles and nucleic acids, products, and antibodies thereto that are involved in cancer and other diseases; and to the use of such expression profiles and compositions in the diagnosis, prognosis, and therapy of these conditions. The invention further relates to methods for identifying and using agents and/or targets that modulate these conditions.

BACKGROUND OF THE INVENTION

Cancer is a major cause of morbidity in the United States. For example, in 1996, the American Cancer Society estimated that 1,359,150 people were diagnosed with a malignant neoplasm and 554,740 died from one of these diseases. Cancer is responsible for 23.9 percent of all American deaths and is exceeded only by heart disease as a cause of mortality (33 percent). Unfortunately, cancer mortality is increasing and sometime early in this century, cancer is expected to become the leading cause of mortality in the United States as it already is in Japan.

Cancers share the charactaristic of disordered control over normal cell division, growth, and differentiation. Their initial clinical manifestations are extremely heterogeneous, with over 70 types of cancer arising in virtually every organ and tissue of the body. Moreover, some of those similarly classified cancer types may represent multiple different molecular diseases. Unfortunately, some cancers may be virtually asymptomatic until late in the disease course, when treatment is more difficult, and prognosis grim.

Treatment for cancer typically includes surgery, chemotherapy, and/or radiation therapy. Although nearly 50 percent of cancer patients can be effectively treated using these methods, the current therapies all induce serious side effects which diminish quality of life. The identification of novel therapeutic targets and diagnostic markers will be important for improving the diagnosis, prognosis, and treatment of cancer patients.

Recent advances in molecular medicine have increased the interest in tumor-specific antigens that could serve as targets for various immunotherapeutic or small molecule strategies. Antigens suitable for immunotherapeutic strategies should be highly expressed in cancer tissues, preferably accessible from the vasculature and at the cell surface, and

ideally not expressed in normal adult tissues. Expression in tissues that are dispensable for life, however, may be tolerated, e.g., reproductive organs, especially those absent in one sex. Examples of antigens that are currently available for the detection and treatment of certain cancers include Her2/neu and the B-cell antigen CD20. Humanized monclonal antibodies directed to Her2/neu (Herceptin®/trastuzumab) are currently in use for the treatment of metastatic breast cancer. See Ross and Fletcher (1998) Stem Cells 16:413-428. Similarly, anti-CD20 monoclonal antibodies (Rituxin®/rituximab) are used to effectively treat non-Hodgkin's lymphoma. See Maloney, et al. (1997) Blood 90:2188-2195; Leget and Czuczman (1998) Curr. Opin. Oncol. 10:548-551.

The elucidation of a role for novel proteins and compounds in disease states for identification of therapeutic targets and diagnostic markers is valuable for improving the current treatment of cancer patients. Accordingly, provided herein are molecular targets for therapeutic intervention in various defined cancers. Additionally, provided herein are methods that can be used in diagnosis and prognosis of cancer. Further provided are methods that can be used to screen candidate bioactive agents for the ability to modulate cancer.

SUMMARY OF THE INVENTION

The present invention provides methods for detecting a pathological cell in a patient, the method comprising detecting a nucleic acid or polypeptide comprising a sequence at least 80% identical to a sequence described in Table 2 or the attached listing of SEQ ID NOs:1-116 in a biological sample from the patient, thereby detecting, either qualitatively or quantitatively, the pathological cell. In certain embodiments of the method, the pathological cell has a pathology (i.e. disease state, abnormality, or medical condition) selected from those listed in Table 1, including cancer. In some embodiments of the method, the biological sample comprises nucleic acids (e.g. mRNA); the biological sample is tissue from an organ which is affected by a pathology listed in Table 1, including a cancer; a further step is used of amplifying nucleic acids before the step of detecting the nucleic acid; the detecting is of a protein encoded by the nucleic acid; the nucleic acid comprises a sequence as described in Table 2 or the attached listing of SEQ ID NOs:1-116 ; the detecting step is carried out by using a labeled nucleic acid probe, utilizing a biochip comprising a sequence at least 80% identical to a sequence as described in Table 2 or the attached listing of SEQ ID NOs:1-116 , or detecting a polypeptide encoded by a nucleic

acid; or the patient is undergoing a therapeutic regimen to treat a pathology of Table 1, or is suspected of having a pathology (e.g. cancer).

Compositions are also provided, e.g., an isolated nucleic acid molecule comprising a sequence as described in Table 2 or SEQ ID NOs:1-58, including, e.g., those which are labeled; an expression vector comprising such nucleic acid; a host cell comprising such expression vector; an isolated polypeptide which is encoded by such a nucleic acid molecule comprising a sequence as described in Table 2 or SEQ ID NOs:59-116; or an antibody that specifically binds a polypeptide comprising a sequence selected from those listed in SEQ ID NOs:59-116. In particular embodiments, the antibody is conjugated to an effector component, is conjugated to a detectable label (including, e.g., a fluorescent label, a radioisotope, or a cytotoxic chemical), an antibody fragment, or is a humanized antibody.

Additional methods are provided, including methods for specifically targeting a compound to a pathological cell in a patient, the method comprising administering to the patient an antibody conjugated to, or capable of binding to, the compound, as described, thereby providing the targetting. Others include, e.g., methods for determining the presence or absence of a pathological cell in a patient, the methods comprising contacting a biological sample with an antibody, as described. In more particular methods, the antibody is: conjugated to an effector component, or to a fluorescent label; or the biological sample is a blood, serum, urine, or stool sample.

Further methods include those for identifying, or screening, compounds that modulate the function of pathology-associated polypeptides (e.g. polypeptides that have been identified associated with a disease state via gene expression analysis), the method comprising: contacting the compound with a pathology-associated polypeptide, the polypeptide encoded by a polynucleotide that selectively hybridizes to a sequence at least 80% identical to a sequence as described in Table 2 or the attached listing of SEQ ID NOs:1-116 ; and determining the effect of the compound upon the function of the polypeptide. Another drug screening assay method comprises steps of: administering a test compound to a mammal having a pathology of Table 1 or a cell isolated therefrom; and comparing the level of gene expression of a polynucleotide that selectively hybridizes to a sequence at least 80% identical to a sequence as described in Table 2 or the attached listing of SEQ ID NOs:1-116 in a treated cell or mammal with the level of gene expression of the polynucleotide in a control cell or mammal, wherein a test compound that modulates the level of expression of the polynucleotide is a candidate for the treatment of the pathology.

DETAILED DESCRIPTION OF THE INVENTION

In accordance with the objects outlined above, the present invention provides novel methods for diagnosis and prognosis evaluation for various disorders, e.g., angiogenesis, fibrosis, and various defined forms of cancer, including metastatic cancer, as well as methods for screening for compositions which modulate such conditions. Also provided are methods for treating such disorders or cancers. See, e.g., American Society of Clinical Oncology (ed. 2001) ASCO Curriculum: Symptom Management Kendall/Hunt, ISBN: 0787277851; Bonadonna, et al. (2001) Textbook of Breast Cancer (2d ed.) Dunitz Martin, ISBN: 1853178241; Devita and Hellman (eds. 2001) Cancer Principles and Practice of Oncology (2 vols.), Lippincott Williams, ISBN: 0781723876; Howell, et al. (2001) Breast Cancer Isis Medical Media, ISBN: 1901865584; Kaye and Laws (2001) Brain Tumours: An Encyclopedic Approach (2d ed.) Churchill Livingstone, ISBN: 0443064261; Mihm, et al. (2001) The Melanocytic Proliferation: A Comprehensive Textbook of Pigmented Lesions Wiley-Liss, ISBN: 0471252719; Montgomery and Aaron (2001) Clinical Pathology of Soft-Tissue Tumors Marcel Dekker, ISBN: 0824702905; Petrovich, et al. (eds. 2001) Combined Modality of Central Nervous System Tumors (Medical Radiology) Springer Verlag, ISBN: 3540660534; Rosen (2001) Rosen's Breast Pathology Lippincott Williams and Wilkins, ISBN: 0781723795; Shah, et al. (2001) Oral Cancer Isis Medical Media, ISBN: 189906687X; Weiss and Goldblum (2001) Enzinger and Weiss's Soft Tissue Tumors (4th ed.) Mosby, ISBN: 0323012000; Abeloff, et al. (eds. 2000) Clinical Oncology (2d ed.) Churchill Livingstone, ISBN: 044307545X; American Society of Clinical Oncology (ed. 2000) Cancer Genetics and Cancer Predisposition Testing Kendall/Hunt, ISBN: 0787276154; Fletcher (2000) Diagnostic Histopathology of Tumors (2 vols. 2d ed.) Churchill Livingstone, ISBN: 0443079927; Vogelzang (ed. 2000) Comprehensive Textbook of Genitourinary Oncology (2d ed.) Lippincott Williams and Wilkins, ISBN: 0683306456; Holland, et al. (eds. 2000) Holland-Frei Cancer Medicine (Book with CD-ROM 5th ed.) Decker, ISBN: 1550091131; Turrisi, et al. (2000) Lung Cancer Isis Medical Media, ISBN: 1901865428; Bartolozzi and Lencioni (eds. 1999) Liver Malignancies: Diagnostic and Interventional Radiology (Medical Radiology) Springer Verlag, ISBN: 3540647562; Gasparini (ed. 1999) Prognostic Variables in Node-Negative and Node-Positive Breast Cancer Kluwer, ISBN: 0792384474; Hansen (ed. 1999) The LASLC Textbook of Lung Cancer: International Association for the Study of Lung Cancer Dunitz Martin, ISBN: 1853177083; Raghavan, et al. (eds. 1999) Textbook of Uncommon Cancer (2nd ed.) Wiley,

ISBN: 0471929212; Thawley, et al. (eds. 1999) Comprehensive Management of Head and Neck Tumors (2 vols.) Saunders, ISBN: 0721655823; Whittaker and Holmes (eds. 1999) Leukemia and Related Disorders (3d ed.) Blackwell Science, ISBN: 0865426074; Aapro (ed. 1998) OncoMedia: Medical Oncology (CD-ROM) Elsevier Science, ISBN: 0080427480; Abeloff (1998) Clinical Oncology (Library Version 2 CD-ROM Individual Version 2.0 Windows and Macintosh) Harcourt Brace, ISBN: 0443075557; Benson (ed. 1998) Gastrointestinal Oncology (Cancer Treatment and Research, CTAR 98) Kluwer, ISBN: 0792382056; Brambilla and Brambilla (eds. 1998) Lung Tumors: Fundamental Biology and Clinical Management (Vol 124) Marcel Dekker, ISBN: 0824701607; Canellos, et al. (eds. 1998) The Lymphomas Saunders, ISBN: 0721650309; Greenspan and Remagen (1998) Differential Diagnosis of Tumors and Tumor-Like Lesions of Bones and Joints Lippincott Williams and Wilkins Publishers, ISBN: 0397517106; Hiddemann (ed. 1998) Acute Leukemias VII: Experimental Approaches and Novel Therapies (Haematologie Und Bluttransfusion, Vol 39), Springer Verlag, ISBN: 3540635041; Husband and Reznek (1998) Imaging in Oncology (2 vols.) Mosby, ISBN: 1899066489; Leibel and Phillips (eds. 1998) Textbook of Radiation Oncology Saunders, ISBN: 0721653367; Maloney and Miller (eds. 1998) Cutaneous Oncology: Pathophysiology, Diagnosis, and Management Blackwell Science, ISBN: 0865425175; Mittal, et al. (eds. 1998) Advances in Radiation Therapy Kluwe, ISBN: 0792399811; Oldham (ed. 1998) Principles of Cancer Biotherapy (3d ed.) Kluwer, ISBN: 0792335074; Ozols (ed. 1998) Gynecologic Oncology Kluwer, ISBN: 0792380703; Parkin, et al. (eds. 1998) Cancer Incidence in Five Continents (Iarc Scientific Publications, No 143) Oxford University Press, ISBN: 9283221435; Perez and Brady (eds. 1998) Principles and Practice of Radiation Oncology Lippincott Williams and Wilkins, ISBN: 0397584164; Black, et al. (eds. 1997) Cancer of the Nervous System Blackwell Science, ISBN: 0865423849; Bonadonna, et al. (1997) Textbook of Breast Cancer: A Clinical Guide to Therapy Blackwell Science, ISBN: 1853173487; Pollock (ed. 1997) Surgical Oncology Kluwer, ISBN: 0792399005; Sheaves, et al. (eds. 1997) Clinical Endocrine Oncology Blackwell Science, ISBN: 086542862X; Vahrson (1997) Radiation Oncology of Gynecological Cancers Springer Verlag, ISBN: 0387567682; Walterhouse and Cohn (eds. 1997) Diagnostic and Therapeutic Advances in Pediatric Oncology Kluwer, ISBN: 0792399781; Aisner (ed. 1996) Comprehensive Textbook of Thoracic Oncology Lippincott, Williams and Wilkins, ISBN: 0683000624; Bertino, et al. (eds. 1996) Encyclopedia of Cancer (3 vols.) Academic, ISBN: 012093230X; Cavalli, et al. (1996)

Textbook of Medical Oncology Dunitz Martin, ISBN: 1853172901; Peckham, et al. (eds. 1995) Oxford Textbook of Oncology (2-Vols.) Oxford University Press, ISBN: 0192616854; and Freireich and Kantarjian (eds. 1996) Molecular Genetics and Therapy of Leukemia (Cancer Treatment and Research, V. 84) Kluwer, ISBN: 0792339126.

In particular, identification of markers selectively expressed on defined cancers allows for use of that expression in diagnostic, prognostic, or therapeutic methods. As such, the invention defines various compositions, e.g., nucleic acids, polypeptides, antibodies, and small molecule agonists/antagonists, which will be useful to selectively identify those markers. For example, therapeutic methods may take the form of protein therapeutics which use the marker expression for selective localization or modulation of function (for those markers which have a causative disease effect), for vaccines, identification of binding partners, or antagonism, e.g., using antisense or RNAi. The markers may be useful for molecular characterization of subsets of the diseases, e.g., as provided in Table 1, which subsets may actually require very different treatments. Moreover, the markers may also be important in related diseases to the specific disorders and cancers, e.g., which affect similar tissues in non-malignant diseases, or have similar mechanisms of induction/maintenance. Metastatic processes or characteristics may also be targeted. Diagnostic and prognostic uses are made available, e.g., to subset related but distinct diseases, or to determine treatment strategy. The detection methods may be based upon nucleic acid, e.g., PCR or hybridization techniques, or protein, e.g., ELISA, imaging, IHC, etc. The diagnosis may be qualitative or quantitative, and may detect increases or decreases in expression levels.

Table 2 provides unigene cluster identification numbers for the nucleotide sequence of genes (SEQ ID NOs:1-58) that exhibit increased or decreased expression in diseased samples, particularly sequences involved in angiogenesis, arthritis, prostate cancer, breast cancer, colorectal cancer, cervical cancer, bladder cancer, head and neck cancer, esophageal cancer, lung cancer, ovarian cancer, pancreatic cancer, renal cancer, stomach cancer, skin cancer, testicular cancer, uterine cancer, glioblastoma, Ewing sarcoma, soft tissue sarcoma, and lung fibrosis. Table 2 also provides an exemplar accession number that provides a nucleotide sequence that is part of the unigene cluster.

Definitions

The term "cancer protein" or "cancer polynucleotide" or "cancer-associated transcript" refers to nucleic acid and polypeptide polymorphic variants, alleles, mutants, and

interspecies homologues that: (1) have a nucleotide sequence that has greater than about 60% nucleotide sequence identity, 65%, 70%, 75%, 80%, 85%, 90%, preferably about 92%, 94%, 96%, 97%, 98%, or 99% or greater nucleotide sequence identity, preferably over a region of over a region of at least about 25, 50, 100, 200, 500, 1000, or more nucleotides, to a nucleotide sequence of or associated with a gene of Table 2 or SEQ ID NOs:1-58; (2) bind to antibodies, e.g., polyclonal antibodies, raised against an immunogen comprising an amino acid sequence encoded by a nucleotide sequence of or associated with a gene of Table 2 or SEQ ID NOs:1-58, and conservatively modified variants thereof; (3) specifically hybridize under stringent hybridization conditions to a nucleic acid sequence, or the complement thereof of Table 2 or SEQ ID NOs:1-58 and conservatively modified variants thereof; or (4) have an amino acid sequence that has greater than about 60% amino acid sequence identity, 65%, 70%, 75%, 80%, 85%, preferably 90%, 91%, 93%, 95%, 97%, 98%, or 99% or greater amino sequence identity, preferably over a region of over a region of at least about 25, 50, 100, 200, 500, 1000, or more amino acids, to an amino acid sequence encoded by a nucleotide sequence of or associated with a gene of Table 2 or SEQ ID NOs:1-58. A polynucleotide or polypeptide sequence is typically from a mammal including, but not limited to, primate, e.g., human; rodent, e.g., rat, mouse, hamster; cow, pig, horse, sheep, or other mammal. A "cancer polypeptide" and a "cancer polynucleotide," include both naturally occurring or recombinant forms.

A "full length" cancer protein or nucleic acid refers to a cancer polypeptide or polynucleotide sequence, or a variant thereof, that contains elements normally contained in one or more naturally occurring, wild type cancer polynucleotide or polypeptide sequences. The "full length" may be prior to, or after, various stages of post-translational processing or splicing, including alternative splicing.

"Biological sample" as used herein is a sample of biological tissue or fluid that contains nucleic acids or polypeptides, e.g., of a cancer protein, polynucleotide, or transcript. Such samples include, but are not limited to, tissue isolated from primates, e.g., humans, or rodents, e.g., mice, and rats. Biological samples may also include sections of tissues such as biopsy and autopsy samples, frozen sections taken for histologic purposes, archival samples, blood, plasma, serum, sputum, stool, tears, mucus, hair, skin, etc. Biological samples also include explants and primary and/or transformed cell cultures derived from patient tissues. A biological sample is typically obtained from a eukaryotic organism, most preferably a mammal such as a primate, e.g., chimpanzee or human; cow;

dog; cat; a rodent, e.g., guinea pig, rat, mouse; rabbit; or a bird; reptile; or fish. Livestock and domestic animals are of interest.

"Providing a biological sample" means to obtain a biological sample for use in methods described in this invention. Most often, this will be done by removing a sample of cells from an animal, but can also be accomplished by using previously isolated cells (e.g., isolated by another person, at another time, and/or for another purpose), or by performing the methods of the invention in vivo. Archival tissues or materials, having treatment or outcome history, will be particularly useful.

The terms "identical" or percent "identity," in the context of two or more nucleic acids or polypeptide sequences, refer to two or more sequences or subsequences that are the same or have a specified percentage of amino acid residues or nucleotides that are the same (e.g., about 70% identity, preferably 75%, 80%, 85%, 90%, 91%, 93%, 95%, 97%, 98%, 99%, or higher identity over a specified region, when compared and aligned for maximum correspondence over a comparison window or designated region) as measured using, e.g., a BLAST or BLAST 2.0 sequence comparison algorithms with default parameters described below, or by manual alignment and visual inspection (see, e.g., the NCBI web site, or the like). Such sequences are then said to be "substantially identical." This definition also refers to, or may be applied to, the complement of a test sequence. The definition also includes sequences that have deletions and/or insertions, substitutions, and naturally occurring, e.g., polymorphic or allelic variants, and man-made variants. As described below, the preferred algorithms can account for gaps and the like. Preferably, identity exists over a region that is at least about 25 amino acids or nucleotides in length, or more preferably over a region that is about 50-100 amino acids or nucleotides in length.

For sequence comparison, typically one sequence acts as a reference sequence, to which test sequences are compared. When using a sequence comparison algorithm, test and reference sequences are entered into a computer, subsequence coordinates are designated, if necessary, and sequence algorithm program parameters are designated. Preferably, default program parameters can be used, or alternative parameters can be designated. The sequence comparison algorithm then calculates the percent sequence identities for the test sequences relative to the reference sequence, based on the program parameters.

A "comparison window", as used herein, includes reference to a segment of contiguous positions selected from the group consisting typically of from about 20 to 600, usually about 50 to 200, more usually about 100 to 150, in which a sequence may be

compared to a reference sequence of the same number of contiguous positions after the two sequences are optimally aligned. Methods of alignment of sequences for comparison are well-known. Optimal alignment of sequences for comparison can be conducted, e.g., by the local homology algorithm of Smith and Waterman (1981) Adv. Appl. Math. 2:482-489, by the homology alignment algorithm of Needleman and Wunsch (1970) J. Mol. Biol. 48:443-453, by the search for similarity method of Pearson and Lipman (1988) Proc. Nat'l. Acad. Sci. USA 85:2444-2448, by computerized implementations of these algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI), or by manual alignment and visual inspection (see, e.g., Ausubel, et al. (eds. 1995 and supplements) Current Protocols in Molecular Biology Wiley).

Preferred examples of algorithms that are suitable for determining percent sequence identity and sequence similarity include the BLAST and BLAST 2.0 algorithms, which are described in Altschul, et al. (1977) Nuc. Acids Res. 25:3389-3402 and Altschul, et al. (1990) J. Mol. Biol. 215:403-410. BLAST and BLAST 2.0 are used, with the parameters described herein, to determine percent sequence identity for the nucleic acids and proteins of the invention. Software for performing BLAST analyses is publicly available through the web-site for National Center for Biotechnology Information (NCBI). This algorithm involves first identifying high scoring sequence pairs (HSPs) by identifying short words of length W in the query sequence, which either match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. T is referred to as the neighborhood word score threshold (Altschul, et al., supra). These initial neighborhood word hits act as seeds for initiating searches to find longer HSPs containing them. The word hits are extended in both directions along each sequence for as far as the cumulative alignment score can be increased. Cumulative scores are calculated using, e.g., for nucleotide sequences, the parameters M (reward score for a pair of matching residues; always > 0) and N (penalty score for mismatching residues; always < 0). For amino acid sequences, a scoring matrix is used to calculate the cumulative score. Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T, and X determine the sensitivity and speed of the alignment. The BLASTN program (for nucleotide sequences)

uses as defaults a wordlength (W) of 11, an expectation (E) of 10, M=5, N=-4 and a comparison of both strands. For amino acid sequences, the BLASTP program uses as defaults a wordlength of 3, and expectation (E) of 10, and the BLOSUM62 scoring matrix (see Henikoff and Henikoff (1992) Proc. Natl. Acad. Sci. USA 89:10915-919) alignments (B) of 50, expectation (E) of 10, M=5, N=-4, and a comparison of both strands.

The BLAST algorithm also performs a statistical analysis of the similarity between two sequences. See, e.g., Karlin and Altschul (1993) Proc. Nat'l. Acad. Sci. USA 90:5873-5787. One measure of similarity provided by the BLAST algorithm is the smallest sum probability (P(N)), which provides an indication of the probability by which a match between two nucleotide or amino acid sequences would occur by chance. For example, a nucleic acid is considered similar to a reference sequence if the smallest sum probability in a comparison of the test nucleic acid to the reference nucleic acid is less than about 0.2, more preferably less than about 0.01, and most preferably less than about 0.001. Log values may be negative large numbers, e.g., 5, 10, 20, 30, 40, 40, 70, 90, 110, 150, 170, etc.

An indication that two nucleic acid sequences are substantially identical is that the polypeptide encoded by the first nucleic acid is immunologically cross reactive with the antibodies raised against the polypeptide encoded by the second nucleic acid. Thus, a polypeptide is typically substantially identical to a second polypeptide, e.g., where the two peptides differ only by conservative substitutions. Another indication that two nucleic acid sequences are substantially identical is that the two molecules or their complements hybridize to each other under stringent conditions. Yet another indication that two nucleic acid sequences are substantially identical is that the same primers can be used to amplify the sequences.

A "host cell" is a naturally occurring cell or a transformed cell that contains an expression vector and supports the replication or expression of the expression vector. Host cells may be cultured cells, explants, cells in vivo, and the like. Host cells may be prokaryotic cells such as E. coli, or eukaryotic cells such as yeast, insect, amphibian, or mammalian cells such as CHO, HeLa, and the like (see, e.g., the American Type Culture Collection (ATCC) catalog or web site).

The terms "isolated," "purified," or "biologically pure" refer to material that is substantially or essentially free from components that normally accompany it as found in its native state. Purity and homogeneity are typically determined using analytical chemistry techniques such as polyacrylamide gel electrophoresis or high performance liquid

chromatography. A protein or nucleic acid that is the predominant species present in a preparation is substantially purified. In particular, an isolated nucleic acid is separated from some open reading frames that naturally flank the gene and encode proteins other than protein encoded by the gene. The term "purified" in some embodiments denotes that a nucleic acid or protein gives rise to essentially one band in an electrophoretic gel. Preferably, it means that the nucleic acid or protein is at least about 85% pure, more preferably at least 95% pure, and most preferably at least 99% pure. "Purify" or "purification" in other embodiments means removing at least one contaminant or component from the composition to be purified. In this sense, purification does not require that the purified compound be homogeneous, e.g., 100% pure.

The terms "polypeptide," "peptide," and "protein" are used interchangeably herein to refer to a polymer of amino acid residues. The terms apply to amino acid polymers in which one or more amino acid residue is an artificial chemical mimetic of a corresponding naturally occurring amino acid, as well as to naturally occurring amino acid polymers, those containing modified residues, and non-naturally occurring amino acid polymers.

The term "amino acid" refers to naturally occurring and synthetic amino acids, as well as amino acid analogs and amino acid mimetics that function similarly to the naturally occurring amino acids. Naturally occurring amino acids are those encoded by the genetic code, as well as those amino acids that are later modified, e.g., hydroxyproline, γ-carboxyglutamate, and O-phosphoserine. Amino acid analogs refers to compounds that have the same basic chemical structure as a naturally occurring amino acid, e.g., an α carbon that is bound to a hydrogen, a carboxyl group, an amino group, and an R group, e.g., homoserine, norleucine, methionine sulfoxide, methionine methyl sulfonium. Such analogs may have modified R groups (e.g., norleucine) or modified peptide backbones, but retain somebasic chemical structure as a naturally occurring amino acid. Amino acid mimetic refers to a chemical compound that has a structure that is different from the general chemical structure of an amino acid, but that functions similarly to another amino acid.

Amino acids may be referred to herein by either their commonly known three letter symbols or by the one-letter symbols recommended by the IUPAC-IUB Biochemical Nomenclature Commission. Nucleotides, likewise, may be referred to by their commonly accepted single-letter codes.

"Conservatively modified variant" applies to both amino acid and nucleic acid sequences. With respect to particular nucleic acid sequences, conservatively modified

variants refers to those nucleic acids which encode identical or essentially identical amino acid sequences, or where the nucleic acid does not encode an amino acid sequence, to essentially identical or associated, e.g., naturally contiguous, sequences. Because of the degeneracy of the genetic code, a large number of functionally identical nucleic acids encode most proteins. For instance, the codons GCA, GCC, GCG, and GCU each encode the amino acid alanine. Thus, at each position where an alanine is specified by a codon, the codon can be altered to another of the corresponding codons described without altering the encoded polypeptide. Such nucleic acid variations are "silent variations," which are one species of conservatively modified variations. Every nucleic acid sequence herein which encodes a polypeptide also describes silent variations of the nucleic acid. In certain contexts each codon in a nucleic acid (except AUG, which is ordinarily the only codon for methionine, and TGG, which is ordinarily the only codon for tryptophan) can be modified to yield a functionally similar molecule. Accordingly, a silent variation of a nucleic acid which encodes a polypeptide is implicit in a described sequence with respect to the expression product, but not necessarily with respect to actual probe sequences.

As to amino acid sequences, one of skill will recognize that individual substitutions, deletions, or additions to a nucleic acid, peptide, polypeptide, or protein sequence which alters, adds, or deletes a single amino acid or a small percentage of amino acids in the encoded sequence is a "conservatively modified variant" where the alteration results in the substitution of an amino acid with a chemically similar amino acid. Conservative substitution tables providing functionally similar amino acids are well known. Such conservatively modified variants are in addition to and do not exclude polymorphic variants, interspecies homologs, and alleles of the invention. Typically conservative substitutions include for one another: 1) Alanine (A), Glycine (G); 2) Aspartic acid (D), Glutamic acid (E); 3) Asparagine (N), Glutamine (Q); 4) Arginine (R), Lysine (K); 5) Isoleucine (I), Leucine (L), Methionine (M), Valine (V); 6) Phenylalanine (F), Tyrosine (Y), Tryptophan (W); 7) Serine (S), Threonine (T); and 8) Cysteine (C), Methionine (M) (see, e.g., Creighton (1984) Proteins: Structure and Molecular Properties Freeman).

Macromolecular structures such as polypeptide structures can be described in terms of various levels of organization. For a general discussion of this organization, see, e.g., Alberts, et al. (eds. 2001) Molecular Biology of the Cell (4th ed.) Garland; and Cantor and Schimmel (1980) Biophysical Chemistry Part I: The Conformation of Biological Macromolecules Freeman. "Primary structure" refers to the amino acid sequence of a

particular peptide. "Secondary structure" refers to locally ordered, three dimensional structures within a polypeptide. These structures are commonly known as domains. Domains are portions of a polypeptide that often form a compact unit of the polypeptide and are typically 25 to approximately 500 amino acids long. Typical domains are made up of sections of lesser organization such as stretches of β-sheet and α-helices. "Tertiary structure" refers to the complete three dimensional structure of a polypeptide monomer. "Quaternary structure" refers to the three dimensional structure formed, usually by the noncovalent association of independent tertiary units. Anisotropic terms are also known as energy terms.

"Nucleic acid" or "oligonucleotide" or "polynucleotide" or grammatical equivalents used herein means at least two nucleotides covalently linked together. Oligonucleotides are typically from about 5, 6, 7, 8, 9, 10, 12, 15, 25, 30, 40, 50, or more nucleotides in length, up to about 100 nucleotides in length. Nucleic acids and polynucleotides are a polymers of any length, including longer lengths, e.g., 200, 300, 500, 1000, 2000, 3000, 5000, 7000, 10,000, etc. A nucleic acid of the present invention will generally contain phosphodiester bonds, although in some cases, nucleic acid analogs are included that may have at least one different linkahge, e.g., phosphoramidate, phosphorothioate, phosphorodithioate, or O-methylphophoroamidite linkages (see Eckstein (1992) Oligonucleotides and Analogues: A Practical Approach Oxford Univ. Press); and peptide nucleic acid backbones and linkages. Other analog nucleic acids include those with positive backbones; non-ionic backbones, and non-ribose backbones, including those described in U.S. Patent Nos. 5,235,033 and 5,034,506, and Chapters 6 and 7 of Sanghvi and Cook (eds. 1994) Carbohydrate Modifications in Antisense Research ACS Symposium Series 580. Nucleic acids containing one or more carbocyclic sugars are also included within one definition of nucleic acids. Modifications of the ribose-phosphate backbone may be done for a variety of reasons, e.g., to increase the stability and half-life of such molecules in physiological environments or as probes on a biochip. Mixtures of naturally occurring nucleic acids and analogs can be made; alternatively, mixtures of different nucleic acid analogs, and mixtures of naturally occurring nucleic acids and analogs may be made.

A variety of references disclose such nucleic acid analogs, including, e.g., phosphoramidate (Beaucage, et al. (1993) Tetrahedron 49:1925-1963 and references therein; Letsinger (1970) J. Org. Chem. 35:3800-3803; Sprinzl, et al. (1977) Eur. J. Biochem. 81:579-589; Letsinger, et al. (1986) Nucl. Acids Res. 14:3487-499; Sawai, et al.

(1984) Chem. Lett. 805, Letsinger, et al. (1988) J. Am. Chem. Soc. 110:4470-4471; and Pauwels, et al. (1986) Chemica Scripta 26:141-149), phosphorothioate (Mag, et al. (1991) Nucleic Acids Res. 19:1437-441; and U.S. Patent No. 5,644,048), phosphorodithioate (Brill, et al. (1989) J. Am. Chem. Soc. 111:2321-2322), O-methylphophoroamidite linkages (see Eckstein (1992) Oligonucleotides and Analogues: A Practical Approach, Oxford Univ. Press), and peptide nucleic acid backbones and linkages (see Egholm (1992) J. Am. Chem. Soc. 114:1895-1897; Meier, et al. (1992) Chem. Int. Ed. Engl. 31:1008-1010; Nielsen (1993) Nature 365:566-568; Carlsson, et al. (1996) Nature 380:207, all of which are incorporated by reference). Other analog nucleic acids include those with positive backbones (Denpcy, et al. (1995) Proc. Natl. Acad. Sci. USA 92:6097-101; non-ionic backbones (U.S. Patent Nos. 5,386,023, 5,637,684, 5,602,240, 5,216,141, and 4,469,863; Kiedrowski, et al. (1991) Angew. Chem. Intl. Ed. English 30:423-426; Letsinger, et al. (1988) J. Am. Chem. Soc. 110:4470-4471; Letsinger, et al. (1994) Nucleoside and Nucleotide 13:1597; Chapters 2 and 3 in Sanghvi and Cook (eds. 1994) Carbohydrate Modifications in Antisense Research ACS Symposium Series 580; Mesmaeker, et al. (1994) Bioorganic and Medicinal Chem. Lett. 4:395-398; Jeffs, et al. (1994) J. Biomolecular NMR 34:17; Horn, et al. (1996) Tetrahedron Lett. 37:743) and non-ribose backbones, including those described in U.S. Patent Nos. 5,235,033 and 5,034,506, and Chapters 6 and 7 in Sanghvi and Cook (eds. 1994) Carbohydrate Modifications in Antisense Research ACS Symposium Series 580. Nucleic acids containing one or more carbocyclic sugars are also included within one definition of nucleic acids (see Jenkins, et al. (1995) Chem. Soc. Rev. pp 169-176). Several nucleic acid analogs are described in Rawls (page 35, June 2, 1997) C&E News.

Particularly preferred are peptide nucleic acids (PNA) which includes peptide nucleic acid analogs. These backbones are substantially non-ionic under neutral conditions, in contrast to the highly charged phosphodiester backbone of naturally occurring nucleic acids. This results in at least two advantages. The PNA backbone exhibits improved hybridization kinetics. PNAs have larger changes in the melting temperature (T_m) for mismatched versus perfectly matched basepairs. DNA and RNA typically exhibit a 2-4° C drop in T_m for an internal mismatch. With the non-ionic PNA backbone, the drop is closer to 7-9° C. Similarly, due to their non-ionic nature, hybridization of the bases attached to these backbones is relatively insensitive to salt concentration. In addition, PNAs are not degraded by cellular enzymes, and thus can be more stable.

The nucleic acids may be single stranded or double stranded, as specified, or contain portions of both double stranded or single stranded sequence. The depiction of a single strand also defines the sequence of the complementary strand; thus the sequences described herein also provide the complement of the sequence. The nucleic acid may be DNA, both genomic and cDNA, RNA, or a hybrid, where the nucleic acid may contain combinations of deoxyribo- and ribo-nucleotides, and combinations of bases, including uracil, adenine, thymine, cytosine, guanine, inosine, xanthine hypoxanthine, isocytosine, isoguanine, etc. "Transcript" typically refers to a naturally occurring RNA, e.g., a pre-mRNA, hnRNA, or mRNA. As used herein, the term "nucleoside" includes nucleotides and nucleoside and nucleotide analogs, and modified nucleosides such as amino modified nucleosides. In addition, "nucleoside" includes non-naturally occurring analog structures. Thus, e.g., the individual units of a peptide nucleic acid, each containing a base, are referred to herein as a nucleoside.

A "label" or a "detectable moiety" is a composition detectable by spectroscopic, photochemical, biochemical, immunochemical, physiological, chemical, or other physical means. In general, labels fall into three classes: a) isotopic labels, which may be radioactive or heavy isotopes; b) immune labels, which may be antibodies, antigens, or epitope tags; and c) colored or fluorescent dyes. The labels may be incorporated into the cancer nucleic acids, proteins, and antibodies. For example, the label should be capable of producing, either directly or indirectly, a detectable signal. The detectable moiety may be a radioisotope, such as ^{3}H, ^{14}C, ^{32}P, ^{35}S, or ^{125}I, electron-dense reagents, a fluorescent or chemiluminescent compound, such as fluorescein isothiocyanate, rhodamine, or luciferin, or an enzyme (e.g., as commonly used in an ELISA), biotin, digoxigenin, or haptens and proteins or other entities which can be made detectable such as alkaline phosphatase, beta-galactosidase, or horseradish peroxidase. Methods are known for conjugating the antibody to the label. See, e.g., Hunter, et al. (1962) Nature 144:945; David, et al. (1974) Biochemistry 13:1014-1021; Pain, et al. (1981) J. Immunol. Meth. 40:219-230; and Nygren (1982) J. Histochem. and Cytochem. 30:407-412.

An "effector" or "effector moiety" or "effector component" is a molecule that is bound (or linked, or conjugated), either covalently, through a linker or a chemical bond, or noncovalently, through ionic, van der Waals, electrostatic, or hydrogen bonds, to an antibody. The "effector" can be a variety of molecules including, e.g., detection moieties including radioactive compounds, fluorescent compounds, enzymes or substrates, tags such

as epitope tags, toxins; activatable moieties, chemotherapeutic agents; lipases; antibiotics; chemoattracting moieties, immune modulators (micA/B), or radioisotopes, e.g., emitting "hard" beta, radiation.

A "labeled nucleic acid probe or oligonucleotide" is one that is bound, e.g., covalently, through a linker or a chemical bond, or noncovalently, through ionic, van der Waals, electrostatic, or hydrogen bonds to a label such that the presence of the probe may be detected by detecting the presence of the label bound to the probe. Alternatively, methods using high affinity interactions may achieve the same results where one of a pair of binding partners binds to the other, e.g., biotin, streptavidin.

As used herein a "nucleic acid probe or oligonucleotide" is a nucleic acid capable of binding to a target nucleic acid of complementary sequence through one or more types of chemical bonds, usually through complementary base pairing, e.g., through hydrogen bond formation. As used herein, a probe may include natural (e.g., A, G, C, or T) or modified bases (7-deazaguanosine, inosine, etc.). In addition, the bases in a probe may be joined by a linkage other than a phosphodiester bond, preferably one that does not functionally interfere with hybridization. Thus, e.g., probes may be peptide nucleic acids in which the constituent bases are joined by peptide bonds rather than phosphodiester linkages. Probes may bind target sequences lacking complete complementarity with the probe sequence depending upon the stringency of the hybridization conditions. The probes are preferably directly labeled, e.g., with isotopes, chromophores, lumiphores, chromogens, or indirectly labeled, e.g., with biotin to which a streptavidin complex may later bind. By assaying for the presence or absence of the probe, one can detect the presence or absence of the select sequence or subsequence. Diagnosis or prognosis may be based at the genomic level, or at the level of RNA or protein expression.

The term "recombinant" when used with reference, e.g., to a cell, or nucleic acid, protein, or vector, indicates that the cell, nucleic acid, protein, or vector, has been modified by the introduction of a heterologous nucleic acid or protein or the alteration of a native nucleic acid or protein, or that the cell is derived from a cell so modified. Thus, e.g., recombinant cells express genes that are not found within the native (non-recombinant) form of the cell or express native genes that are otherwise abnormally expressed, under expressed, or not expressed at all. By the term "recombinant nucleic acid" herein is meant nucleic acid, originally formed in vitro, in general, by the manipulation of nucleic acid, e.g., using polymerases and endonucleases, in a form not normally found in nature. In this

manner, operably linkage of different sequences is achieved. Thus an isolated nucleic acid, in a linear form, or an expression vector formed in vitro by ligating DNA molecules that are not normally joined, are both considered recombinant for the purposes of this invention. It is understood that once a recombinant nucleic acid is made and reintroduced into a host cell or organism, it will replicate non-recombinantly, e.g., using the in vivo cellular machinery of the host cell rather than in vitro manipulations; however, such nucleic acids, once produced recombinantly, although subsequently replicated non-recombinantly, are still considered recombinant for the purposes of the invention.

Similarly, a "recombinant protein" is a protein made using recombinant techniques, e.g., through the expression of a recombinant nucleic acid as depicted above. A recombinant protein is distinguished from naturally occurring protein by at least one or more characteristics. The protein may be isolated or purified away from some or most of the proteins and compounds with which it is normally associated in its wild type host, and thus may be substantially pure. An isolated protein is unaccompanied by at least some of the material with which it is normally associated in its natural state, preferably constituting at least about 0.5%, more preferably at least about 5% by weight of the total protein in a given sample. A substantially pure protein comprises at least about 75% by weight of the total protein, with at least about 80% being preferred, and at least about 90% being particularly preferred. The definition includes the production of a cancer protein from one organism in a different organism or host cell. Alternatively, the protein may be made at a significantly higher concentration than is normally seen, through the use of an inducible promoter or high expression promoter, such that the protein is made at increased concentration levels. Alternatively, the protein may be in a form not normally found in nature, as in the addition of an epitope tag or amino acid substitutions, insertions and deletions, as discussed below.

The term "heterologous" when used with reference to portions of a nucleic acid indicates that the nucleic acid comprises two or more subsequences that are not normally found in the same relationship to each other in nature. For instance, the nucleic acid is typically recombinantly produced, having two or more sequences, e.g., from unrelated genes arranged to make a new functional nucleic acid, e.g., a promoter from one source and a coding region from another source. Similarly, a heterologous protein will often refer to two or more subsequences that are not found in the same relationship to each other in nature (e.g., a fusion protein).

A "promoter" is typically an array of nucleic acid control sequences that direct transcription of a nucleic acid. As used herein, a promoter includes necessary nucleic acid sequences near the start site of transcription, such as, in the case of a polymerase II type promoter, a TATA element. A promoter also optionally includes distal enhancer or repressor elements, which can be located as much as several thousand base pairs from the start site of transcription. A "constitutive" promoter is a promoter that is active under most environmental and developmental conditions. An "inducible" promoter is active under environmental or developmental regulation. The term "operably linked" refers to a functional linkage between a nucleic acid expression control sequence (such as a promoter, or array of transcription factor binding sites) and a second nucleic acid sequence, e.g., wherein the expression control sequence directs transcription of the nucleic acid corresponding to the second sequence.

An "expression vector" is a nucleic acid construct, generated recombinantly or synthetically, with a series of specified nucleic acid elements that permit transcription of a particular nucleic acid in a host cell. The expression vector can be part of a plasmid, virus, or nucleic acid fragment. Typically, the expression vector includes a nucleic acid to be transcribed in operable linkage to a promoter.

The phrase "selectively (or specifically) hybridizes to" refers to the binding, duplexing, or hybridizing of a molecule selectively to a particular nucleotide sequence under stringent hybridization conditions when that sequence is present in a complex mixture (e.g., total cellular or library DNA or RNA).

The phrase "stringent hybridization conditions" refers to conditions under which a probe will hybridize to its target subsequence, typically in a complex mixture of nucleic acids, but to no other sequences. Stringent conditions are sequence-dependent and will be different in different circumstances. Longer sequences hybridize specifically at higher temperatures. An extensive guide to the hybridization of nucleic acids is found in "Overview of principles of hybridization and the strategy of nucleic acid assays" in Tijssen (1993) Hybridization with Nucleic Probes (Laboratory Techniques in Biochemistry and Molecular Biology) (vol. 24) Elsevier. Generally, stringent conditions are selected to be about 5-10° C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength pH. The T_m is the temperature (under defined ionic strength, pH, and nucleic concentration) at which 50% of the probes complementary to the target hybridize to the target sequence at equilibrium (as the target sequences are present in

excess, at T_m, 50% of the probes are occupied at equilibrium). Stringent conditions will be those in which the salt concentration is less than about 1.0 M sodium ion, typically about 0.01-1.0 M sodium ion concentration (or other salts) at pH 7.0 to 8.3 and the temperature is at least about 30° C for short probes (e.g., about 10-50 nucleotides) and at least about 60° C for long probes (e.g., greater than about 50 nucleotides). Stringent conditions may also be achieved with the addition of destabilizing agents such as formamide. For selective or specific hybridization, a positive signal is typically at least two times background, preferably 10 times background hybridization. Exemplary stringent hybridization conditions can be as following: 50% formamide, 5x SSC, and 1% SDS, incubating at 42° C, or, 5x SSC, 1% SDS, incubating at 65° C, with wash in 0.2x SSC, and 0.1% SDS at 65° C. For PCR, a temperature of about 36° C is typical for low stringency amplification, although annealing temperatures may vary between about 32-48° C depending on primer length. For high stringency PCR amplification, a temperature of about 62° C is typical, although high stringency annealing temperatures can range from about 50-65° C, depending on the primer length and specificity. Typical cycle conditions for both high and low stringency amplifications include a denaturation phase of 90-95° C for 30-120 sec, an annealing phase lasting 30-120 sec, and an extension phase of about 72° C for 1-2 min. Protocols and guidelines for low and high stringency amplification reactions are provided, e.g., in Innis, et al. (1990) PCR Protocols: A Guide to Methods and Applications Academic Press, NY.

Nucleic acids that do not hybridize to each other under stringent conditions are still substantially identical if the polypeptides which they encode are substantially identical. This occurs, e.g., when a copy of a nucleic acid is created using the maximum codon degeneracy permitted by the genetic code. In such cases, the nucleic acids typically hybridize under moderately stringent hybridization conditions. Exemplary "moderately stringent hybridization conditions" include a hybridization in a buffer of 40% formamide, 1 M NaCl, 1% SDS at 37° C, and a wash in 1X SSC at 45° C. A positive hybridization is typically at least twice background. Alternative hybridization and wash conditions can be utilized to provide conditions of similar stringency. Additional guidelines for determining hybridization parameters are provided in numerous references, e.g., Ausubel, et al. (eds. 1991 and supplements) Current Protocols in Molecular Biology Wiley.

The phrase "functional effects" in the context of assays for testing compounds that modulate activity of a cancer protein includes the determination of a parameter that is

indirectly or directly under the influence of the cancer protein or nucleic acid, e.g., a physiological, functional, physical, or chemical effect, such as the ability to decrease cancer. It includes ligand binding activity; cell viability; cell growth on soft agar; anchorage dependence; contact inhibition and density limitation of growth; cellular proliferation; cellular transformation; growth factor or serum dependence; tumor specific marker levels; invasiveness into Matrigel; tumor growth and metastasis in vivo; mRNA and protein expression in cells undergoing metastasis; and other characteristics of cancer cells. "Functional effects" include in vitro, in vivo, and ex vivo activities.

By "determining the functional effect" is meant assaying for a compound that increases or decreases a parameter that is indirectly or directly under the influence of a cancer protein sequence, e.g., physiological, functional, enzymatic, physical, or chemical effects. Such functional effects can be measured, e.g., changes in spectroscopic characteristics (e.g., fluorescence, absorbance, refractive index), hydrodynamic (e.g., shape), chromatographic, or solubility properties for the protein, measuring inducible markers or transcriptional activation of the cancer protein, measuring binding activity or binding assays, e.g., binding to antibodies or other ligands, and measuring growth, cellular proliferation, cell viability, cellular transformation, growth factor or serum dependence, tumor specific marker levels, invasiveness into Matrigel, tumor growth and metastasis in vivo, mRNA and protein expression, and other characteristics of cancer cells. The functional effects can be evaluated by many means, e.g., microscopy for quantitative or qualitative measures of alterations in morphological features, measurement of changes in RNA or protein levels for cancer-associated sequences, measurement of RNA stability, identification of downstream or reporter gene expression (CAT, luciferase, β-gal, GFP, and the like), e.g., via chemiluminescence, fluorescence, colorimetric reactions, antibody binding, inducible markers, and ligand binding assays.

"Inhibitors", "activators," and "modulators" of cancer polynucleotide and polypeptide sequences are used to refer to activating, inhibitory, or modulating molecules or compounds identified using in vitro and in vivo assays of cancer polynucleotide and polypeptide sequences. Inhibitors are compounds that, e.g., bind to, partially or totally block activity, decrease, prevent, delay activation, inactivate, desensitize, or down regulate the activity or expression of cancer proteins, e.g., antagonists. Antisense or inhibitory nucleic acids may seem to inhibit expression and subsequent function of the protein. "Activators" are compounds that increase, open, activate, facilitate, enhance activation,

sensitize, agonize, or up regulate cancer protein activity. Inhibitors, activators, or modulators also include genetically modified versions of cancer proteins, e.g., versions with altered activity, as well as naturally occurring and synthetic ligands, antagonists, agonists, antibodies, small chemical molecules, and the like. Such assays for inhibitors and activators include, e.g., expressing the cancer protein in vitro, in cells, or cell membranes, applying putative modulator compounds, and then determining the functional effects on activity, as described above. Activators and inhibitors of cancer can also be identified by incubating cancer cells with the test compound and determining increases or decreases in the expression of 1 or more cancer proteins, e.g., 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 40, 50, or more cancer proteins, such as cancer proteins encoded by the sequences set out in Table 2 or SEQ ID NOs:59-116.

Samples or assays comprising cancer proteins that are treated with a potential activator, inhibitor, or modulator are compared to control samples without the inhibitor, activator, or modulator to examine the extent of inhibition. Control samples (untreated with inhibitors) are assigned a relative protein activity value of 100%. Inhibition of a polypeptide is achieved when the activity value relative to the control is about 80%, preferably 50%, more preferably 25-0%. Activation of a cancer polypeptide is achieved when the activity value relative to the control (untreated with activators) is about 110%, more preferably 150%, more preferably 200-500% (e.g., two to five fold higher relative to the control), more preferably 1000-3000% higher.

The phrase "changes in cell growth" refers to any change in cell growth and proliferation characteristics in vitro or in vivo, such as cell viability, formation of foci, anchorage independence, semi-solid or soft agar growth, changes in contact inhibition and density limitation of growth, loss of growth factor or serum requirements, changes in cell morphology, gaining or losing immortalization, gaining or losing tumor specific markers, ability to form or suppress tumors when injected into suitable animal hosts, and/or immortalization of the cell. See, e.g., pp. 231-241 in Freshney (1994) Culture of Animal Cells a Manual of Basic Technique (2d ed.) Wiley-Liss.

"Tumor cell" refers to precancerous, cancerous, and normal cells in a tumor.

"Cancer cells," "transformed" cells or "transformation" in tissue culture, refers to spontaneous or induced phenotypic changes that do not necessarily involve the uptake of new genetic material. Although transformation can arise from infection with a transforming virus and incorporation of new genomic DNA, or uptake of exogenous DNA, it can also

arise spontaneously or following exposure to a carcinogen, thereby mutating an endogenous gene. Transformation is associated with phenotypic changes, such as immortalization of cells, aberrant growth control, nonmorphological changes, and/or malignancy. See, Freshney (2000) Culture of Animal Cells: A Manual of Basic Technique (4th ed.) Wiley-Liss.

"Antibody" refers to a polypeptide comprising a framework region from an immunoglobulin gene or fragments thereof that specifically binds and recognizes an antigen. The recognized immunoglobulin genes include the kappa, lambda, alpha, gamma, delta, epsilon, and mu constant region genes, as well as the myriad immunoglobulin variable region genes. Light chains are classified as either kappa or lambda. Heavy chains are classified as gamma, mu, alpha, delta, or epsilon, which in turn define the immunoglobulin classes, IgG, IgM, IgA, IgD, and IgE, respectively. Typically, the antigen-binding region of an antibody or its functional equivalent will be most critical in specificity and affinity of binding. See Paul (ed. 1999) Fundamental Immunology (4th ed.) Raven.

An exemplary immunoglobulin (antibody) structural unit comprises a tetramer. Each tetramer is composed of two identical pairs of polypeptide chains, each pair having one "light" (about 25 kD) and one "heavy" chain (about 50-70 kD). The N-terminus of each chain defines a variable region of about 100 to 110 or more amino acids primarily responsible for antigen recognition. The terms variable light chain (V_L) and variable heavy chain (V_H) refer to these light and heavy chains respectively.

Antibodies exist, e.g., as intact immunoglobulins or as a number of well-characterized fragments produced by digestion with various peptidases. Thus, e.g., pepsin digests an antibody below the disulfide linkages in the hinge region to produce $F(ab)'_2$, a dimer of Fab which itself is a light chain joined to V_H-C_H1 by a disulfide bond. The $F(ab)'_2$ may be reduced under mild conditions to break the disulfide linkage in the hinge region, thereby converting the $F(ab)'_2$ dimer into an Fab' monomer. The Fab' monomer is essentially Fab with part of the hinge region (see Paul (ed. 1999) Fundamental Immunology (4th ed.) Raven. While various antibody fragments are defined in terms of the digestion of an intact antibody, one of skill will appreciate that such fragments may be synthesized de novo either chemically or by using recombinant DNA methodology. Thus, the term antibody, as used herein, also includes antibody fragments either produced by the modification of whole antibodies, or those synthesized de novo using recombinant DNA

methodologies (e.g., single chain Fv) or those identified using phage display libraries (see, e.g., McCafferty, et al. (1990) Nature 348:552-554).

For preparation of antibodies, e.g., recombinant, monoclonal, or polyclonal antibodies, many techniques known. See, e.g., Kohler and Milstein (1975) Nature 256:495-497; Kozbor, et al. (1983) Immunology Today 4:72; Cole, et al. (1985) pp. 77-96 in Reisfeld and Sell (1985) Monoclonal Antibodies and Cancer Therapy Liss; Coligan (1991) Current Protocols in Immunology Lippincott; Harlow and Lane (1988) Antibodies: A Laboratory Manual CSH Press; and Goding (1986) Monoclonal Antibodies: Principles and Practice (2d ed.) Academic Press. Techniques for the production of single chain antibodies (U.S. Patent 4,946,778) can be adapted to produce antibodies to polypeptides of this invention. Also, transgenic mice, or other organisms such as other mammals, may be used to express humanized antibodies. Alternatively, phage display technology can be used to identify antibodies and heteromeric Fab fragments that specifically bind to selected antigens. See, e.g., McCafferty, et al. (1990) Nature 348:552-554; Marks, et al. (1992) Biotechnology 10:779-783.

A "chimeric antibody" is an antibody molecule in which (a) the constant region, or a portion thereof, is altered, replaced, or exchanged so that the antigen binding site (variable region) is linked to a constant region of a different or altered class, and/or species, or an entirely different molecule which confers new properties to the chimeric antibody, e.g., an enzyme, toxin, hormone, growth factor, drug, effector function, chemoattractant, immune modulator, etc.; or (b) the variable region, or a portion thereof, is altered, replaced, or exchanged with a variable region having a different or altered antigen specificity.

Identification of cancer-associated sequences

In one aspect, the expression levels of genes are determined in different patient samples for which diagnosis information is desired, to provide expression profiles. An expression profile of a particular sample is essentially a "fingerprint" of the state of the sample; while two states may have any particular gene similarly expressed, the evaluation of a number of genes simultaneously allows the generation of a gene expression profile that is characteristic of the state of the cell. That is, normal tissue may be distinguished from cancerous or metastatic cancerous tissue, or cancer tissue or metastatic cancerous tissue can be compared with tissue from surviving cancer patients. By comparing expression profiles of tissue in known different cancer states, information regarding which genes are important (including both up-and down-regulation of genes) in each of these states is obtained.

Molecular profiling may distinguish subtypes of a currently collective disease designation, e.g., different forms of a cancer.

The identification of sequences that are differentially expressed in cancer versus non-cancer tissue allows the use of this information in a number of ways. For example, a particular treatment regime may be evaluated: does a chemotherapeutic drug act to down-regulate cancer, and thus tumor growth or recurrence, in a particular patient. Alternatively, a treatment step may induce other markers which may be used as targets to destroy tumor cells. Similarly, diagnosis and treatment outcomes may be done or confirmed by comparing patient samples with the known expression profiles. Maliganant disease may be compared to non-malignant conditions. Metastatic tissue can also be analyzed to determine the stage of cancer in the tissue, or origin of primary tumor, e.g., metastasis from a remote primary site. Furthermore, these gene expression profiles (or individual genes) allow screening of drug candidates with an eye to mimicking or altering a particular expression profile; e.g., screening can be done for drugs that suppress the cancer expression profile. This may be done by making biochips comprising sets of the important cancer genes, which can then be used in these screens. These methods can also be done on the protein basis; that is, protein expression levels of the cancer proteins can be evaluated for diagnostic purposes or to screen candidate agents. In addition, the cancer nucleic acid sequences can be administered for gene therapy purposes, including the administration of antisense nucleic acids, or the cancer proteins (including antibodies and other modulators thereof) administered as therapeutic drugs.

Thus the present invention provides nucleic acid and protein sequences that are differentially expressed in cancer relative to normal tissues and/or non-malignant disease, or in different types of related diseases, herein termed "cancer sequences." As outlined below, cancer sequences include those that are up-regulated (e.g., expressed at a higher level) in cancer, as well as those that are down-regulated (e.g., expressed at a lower level). In a preferred embodiment, the cancer sequences are from humans; however, cancer sequences from other organisms may be useful in animal models of disease and drug evaluation; thus, other cancer sequences are provided, from vertebrates, including mammals, including rodents (rats, mice, hamsters, guinea pigs, etc.), primates, farm animals (including sheep, goats, pigs, cows, horses, etc.) and pets (e.g., dogs, cats, etc.). Cancer sequences from other organisms may be obtained using the techniques outlined below.

Cancer sequences can include both nucleic acid and amino acid sequences. In a preferred embodiment, the skin cancer sequences are recombinant nucleic acids. These nucleic acid sequences are useful in a variety of applications, including diagnostic applications, which will detect naturally occurring nucleic acids, as well as screening applications; e.g., biochips comprising nucleic acid probes or PCR microtiter plates with selected probes to the cancer sequences.

A cancer sequence can be initially identified by substantial nucleic acid and/or amino acid sequence homology to the cancer sequences outlined herein. Such homology can be based upon the overall nucleic acid or amino acid sequence, and is generally determined as outlined below, e.g., using homology programs or hybridization conditions.

For identifying cancer-associated sequences, the cancer screen typically includes comparing genes identified in different tissues, e.g., normal and cancerous tissues, cancer and non-malignant conditions, non-malignant conditions and normal tissues, or tumor tissue samples from patients who have metastatic disease vs. non metastatic tissue. Other suitable tissue comparisons include comparing cancer samples with metastatic cancer samples from other cancers, such as lung, stomach, gastrointestinal cancers, etc. Samples of different stages of cancer, e.g., survivor tissue, drug resistant states, and tissue undergoing metastasis, are applied to biochips comprising nucleic acid probes. The samples are first microdissected, if applicable, and treated for preparation of mRNA. Suitable biochips are commercially available, e.g., from Affymetrix, Santa Clara, CA. Gene expression profiles as described herein are generated and the data analyzed.

In one embodiment, the genes showing changes in expression as between normal and disease states are compared to genes expressed in other normal tissues, including, and not limited to lung, heart, brain, liver, stomach, kidney, muscle, colon, small intestine, large intestine, spleen, bone, and/or placenta. In a preferred embodiment, those genes identified during the cancer screen that are expressed in a significant amount in other tissues (e.g., essential organs) are removed from the profile, although in some embodiments, this is not necessary (e.g., where organs may be dispensible, e.g., female or male specific). That is, when screening for drugs, it is usually preferable that the target expression be disease specific, to minimize possible side effects on other organs were there expression.

In a preferred embodiment, cancer sequences are those that are up-regulated in cancer; that is, the expression of these genes is higher in the cancer tissue as compared to non-cancer or non-malignant tissue. "Up-regulation" as used herein often means at least

about a two-fold change, preferably at least about a three fold change, with at least about five-fold or higher being preferred. Another embodiment is directed to sequences up-regulated in non-malignant conditions relative to normal. Uniformity among relevant samples is also preferred.

Unigene cluster identification numbers and accession numbers herein are for the GenBank sequence database and the sequences of the accession numbers are hereby expressly incorporated by reference. GenBank is available, see, e.g., Benson, et al. (1998) Nuc. Acids Res. 26:1-7. Sequences are also available in other databases, e.g., European Molecular Biology Laboratory (EMBL) and DNA Database of Japan (DDBJ). In some situations, the sequences may be derived from assembly of available sequences or be predicted from genomic DNA using exon prediction algorithms, such as FGENESH. See Salamov and Solovyev (2000) Genome Res. 10:516-522. In other situations, sequences have been derived from cloning and sequencing of isolated nucleic acids.

In another preferred embodiment, cancer sequences are those that are down-regulated in the cancer; that is, the expression of these genes is lower in cancer tissue as compared to non-cancerous tissue. "Down-regulation" as used herein often means at least about a two-fold change, preferably at least about a three fold change, with at least about five-fold or higher being preferred.

Informatics

The ability to identify genes that are over or under expressed in cancer can additionally provide high-resolution, high-sensitivity datasets which can be used in the areas of diagnostics, therapeutics, drug development, pharmacogenetics, protein structure, biosensor development, and other related areas. For example, the expression profiles can be used in diagnostic or prognostic evaluation of patients with cancer or related diseases. See Tables 1-2. Or as another example, subcellular toxicological information can be generated to better direct drug structure and activity correlation (see Anderson (June 11-12, 1998) Pharmaceutical Proteomics: Targets, Mechanism, and Function, paper presented at the IBC Proteomics conference, Coronado, CA). Subcellular toxicological information can also be utilized in a biological sensor device to predict the likely toxicological effect of chemical exposures and likely tolerable exposure thresholds (see U.S. Patent No. 5,811,231). Similar advantages accrue from datasets relevant to other biomolecules and bioactive agents (e.g., nucleic acids, saccharides, lipids, drugs, and the like).

Thus, in another embodiment, the present invention provides a database that includes at least one set of assay data. The data contained in the database is acquired, e.g., using array analysis either singly or in a library format. The database can be in a form in which data can be maintained and transmitted, but is preferably an electronic database. The electronic database of the invention can be maintained on any electronic device allowing for the storage of and access to the database, such as a personal computer, but is preferably distributed on a wide area network, such as the World Wide Web.

The focus of the present section on databases that include peptide sequence data is for clarity of illustration only. Similar databases can be assembled for assay data acquired using an assay of the invention.

The compositions and methods for identifying and/or quantitating the relative and/or absolute abundance of a variety of molecular and macromolecular species from a biological sample representing cancer, e.g., the identification of cancer-associated sequences described herein, provide an abundance of information which can be correlated with pathological conditions, predisposition to disease, drug testing, therapeutic monitoring, gene-disease causal linkages, identification of correlates of immunity and physiological status, among others. Although the data generated from the assays of the invention is suited for manual review and analysis, in a preferred embodiment, data processing using high-speed computers is utilized.

An array of methods for indexing and retrieving biomolecular information is available. For example, U.S. Patents 6,023,659 and 5,966,712 disclose a relational database system for storing biomolecular sequence information in a manner that allows sequences to be catalogued and searched according to one or more protein function hierarchies. U.S. Patent 5,953,727 discloses a relational database having sequence records containing information in a format that allows a collection of partial-length DNA sequences to be catalogued and searched according to association with one or more sequencing projects for obtaining full-length sequences from the collection of partial length sequences. U.S. Patent 5,706,498 discloses a gene database retrieval system for making a retrieval of a gene sequence similar to a sequence data item in a gene database based on the degree of similarity between a key sequence and a target sequence. U.S. Patent 5,538,897 discloses a method using mass spectroscopy fragmentation patterns of peptides to identify amino acid sequences in computer databases by comparison of predicted mass spectra with experimentally-derived mass spectra using a closeness-of-fit measure. U.S. Patent

5,926,818 discloses a multi-dimensional database comprising a functionality for multi-dimensional data analysis described as on-line analytical processing (OLAP), which entails the consolidation of projected and actual data according to more than one consolidation path or dimension. U.S. Patent 5,295,261 reports a hybrid database structure in which the fields of each database record are divided into two classes, navigational and informational data, with navigational fields stored in a hierarchical topological map which can be viewed as a tree structure or as the merger of two or more such tree structures. See also Baxevanis, et al. (2001) Bioinformatics: A Practical Guuide to the Analysis of Genes and Proteins Wiley; Mount (2001) Bioinformatics: Sequence and Genome Analysis CSH Press, NY; Durbin, et al. (eds. 1999) Biological Sequence Analysis: Probabilistic Models of Proteins and Nucleic Acids Cambridge University Press; Baxevanis and Oeullette (eds. 1998) Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins (2d. ed.) Wiley-Liss; Rashidi and Buehler (1999) Bioinformatics: Basic Applications in Biological Science and Medicine CRC Press; Setubal, et al. (eds. 1997) Introduction to Computational Molecular Biology Brooks/Cole; Misener and Krawetz (eds. 2000) Bioinformatics: Methods and Protocols Humana Press; Higgins and Taylor (eds. 2000) Bioinformatics: Sequence, Structure, and Databanks: A Practical Approach Oxford University Press; Brown (2001) Bioinformatics: A Biologist's Guide to Biocomputing and the Internet Eaton Pub.; Han and Kamber (2000) Data Mining: Concepts and Techniques Kaufmann Pub.; and Waterman (1995) Introduction to Computational Biology: Maps, Sequences, and Genomes Chap and Hall.

The present invention provides a computer database comprising a computer and software for storing in computer-retrievable form assay data records cross-tabulated, e.g., with data specifying the source of the target-containing sample from which each sequence specificity record was obtained.

In an exemplary embodiment, at least one of the sources of target-containing sample is from a control tissue sample known to be free of pathological disorders. In a variation, at least one of the sources is a known pathological tissue specimen, e.g., a neoplastic lesion or another tissue specimen to be analyzed for cancer. In another variation, the assay records cross-tabulate one or more of the following parameters for each target species in a sample: (1) a unique identification code, which can include, e.g., a target molecular structure and/or characteristic separation coordinate (e.g., electrophoretic coordinates); (2) sample source; and (3) absolute and/or relative quantity of the target species present in the sample.

The invention also provides for the storage and retrieval of a collection of target data in a computer data storage apparatus, which can include magnetic disks, optical disks, magneto-optical disks, DRAM, SRAM, SGRAM, SDRAM, RDRAM, DDR RAM, magnetic bubble memory devices, and other data storage devices, including CPU registers and on-CPU data storage arrays. Typically, the target data records are stored as a bit pattern in an array of magnetic domains on a magnetizable medium or as an array of charge states or transistor gate states, such as an array of cells in a DRAM device (e.g., each cell comprised of a transistor and a charge storage area, which may be on the transistor). In one embodiment, the invention provides such storage devices, and computer systems built therewith, comprising a bit pattern encoding a protein expression fingerprint record comprising unique identifiers for at least 10 target data records cross-tabulated with target source.

When the target is a peptide or nucleic acid, the invention preferably provides a method for identifying related peptide or nucleic acid sequences, comprising performing a computerized comparison between a peptide or nucleic acid sequence assay record stored in or retrieved from a computer storage device or database and at least one other sequence. The comparison can include a sequence analysis or comparison algorithm or computer program embodiment thereof (e.g., FASTA, TFASTA, GAP, BESTFIT) and/or the comparison may be of the relative amount of a peptide or nucleic acid sequence in a pool of sequences determined from a polypeptide or nucleic acid sample of a specimen.

The invention also preferably provides a magnetic disk, such as an IBM-compatible (DOS, Windows, Windows95/98/2000, Windows NT, OS/2) or other format (e.g., Linux, SunOS, Solaris, AIX, SCO Unix, VMS, MV, Macintosh, etc.) floppy diskette or hard (fixed, Winchester) disk drive, comprising a bit pattern encoding data from an assay of the invention in a file format suitable for retrieval and processing in a computerized sequence analysis, comparison, or relative quantitation method.

The invention also provides a network, comprising a plurality of computing devices linked via a data link, such as an Ethernet cable (coax or 10BaseT), telephone line, ISDN line, wireless network, optical fiber, or other suitable signal transmission medium, whereby at least one network device (e.g., computer, disk array, etc.) comprises a pattern of magnetic domains (e.g., magnetic disk) and/or charge domains (e.g., an array of DRAM cells) composing a bit pattern encoding data acquired from an assay of the invention.

The invention also provides a method for transmitting assay data that includes generating an electronic signal on an electronic communications device, such as a modem, ISDN terminal adapter, DSL, cable modem, ATM switch, or the like, wherein the signal includes (in native or encrypted format) a bit pattern encoding data from an assay or a database comprising a plurality of assay results obtained by the method of the invention.

In a preferred embodiment, the invention provides a computer system for comparing a query target to a database containing an array of data structures, such as an assay result obtained by the method of the invention, and ranking database targets based on the degree of identity and gap weight to the target data. A central processor is preferably initialized to load and execute the computer program for alignment and/or comparison of the assay results. Data for a query target is entered into the central processor via an I/O device. Execution of the computer program results in the central processor retrieving the assay data from the data file, which comprises a binary description of an assay result.

The target data or record and the computer program can be transferred to secondary memory, which is typically random access memory (e.g., DRAM, SRAM, SGRAM, or SDRAM). Targets are ranked according to the degree of correspondence between a selected assay characteristic (e.g., binding to a selected affinity moiety) and the same characteristic of the query target and results are output via an I/O device. For example, a central processor can be a conventional computer (e.g., Intel Pentium, PowerPC, Alpha, PA-8000, SPARC, MIPS 4400, MIPS 10000, VAX, etc.); a program can be a commercial or public domain molecular biology software package (e.g., UWGCG Sequence Analysis Software, Darwin); a data file can be an optical or magnetic disk, a data server, a memory device (e.g., DRAM, SRAM, SGRAM, SDRAM, EPROM, bubble memory, flash memory, etc.); an I/O device can be a terminal comprising a video display and a keyboard, a modem, an ISDN terminal adapter, an Ethernet port, a punched card reader, a magnetic strip reader, or other suitable I/O device.

The invention also preferably provides the use of a computer system, such as that described above, which comprises: (1) a computer; (2) a stored bit pattern encoding a collection of peptide sequence specificity records obtained by the methods of the invention, which may be stored in the computer; (3) a comparison target, such as a query target; and (4) a program for alignment and comparison, typically with rank-ordering of comparison results on the basis of computed similarity values. See, e.g., Ewens and Grant (2001) Statistical Methods in Bioinformatics: An Introduction Springer-Verlag. Mathematical

approaches can also be used to conclude whether similarities or differences in the gene expression exhibited by different samples are significant. See, e.g., Golub, et al. (1999) Science 286:531-537; Duda, et al. (2001) Pattern Classification Wiley; and Hastie, et al. (2001) The Elements of Statistical Learning: Data Mining, Inference, and Prediction Springer-Verlag. One approach to determine whether a sample is more similar to or has maximum similarity with a given condition between the sample and one or more pools representing different conditions for comparison; the pool with the smallest vector angle is then chosen as the most similar to the biological sample among the pools compared.

Characteristics of cancer-associated proteins

Cancer proteins of the present invention may be classified as secreted proteins, transmembrane proteins, or intracellular proteins. In one embodiment, the cancer protein is an intracellular protein. Intracellular proteins may be found in the cytoplasm and/or in the nucleus. Intracellular proteins are involved in all aspects of cellular function and replication (including, e.g., signaling pathways); aberrant expression of such proteins often results in unregulated or disregulated cellular processes (see, e.g., Alberts, et al. (eds. 1994) Molecular Biology of the Cell (3d ed.) Garland). For example, many intracellular proteins have enzymatic activity such as protein kinase activity, protein phosphatase activity, protease activity, nucleotide cyclase activity, polymerase activity, and the like. Intracellular proteins also serve as docking proteins that are involved in organizing complexes of proteins, or targeting proteins to various subcellular localizations, and are involved in maintaining the structural integrity of organelles.

An increasingly appreciated concept in characterizing proteins is the presence in the proteins of one or more structural motifs for which defined functions have been attributed. In addition to the highly conserved sequences found in the enzymatic domain of proteins, highly conserved sequences have been identified in proteins that are involved in protein-protein interaction. For example, Src-homology-2 (SH2) domains bind tyrosine-phosphorylated targets in a sequence dependent manner. PTB domains, which are distinct from SH2 domains, also bind tyrosine phosphorylated targets. SH3 domains bind to proline-rich targets. In addition, PH domains, tetratricopeptide repeats and WD domains to name only a few, have been shown to mediate protein-protein interactions. Some of these may also be involved in binding to phospholipids or other second messengers. These motifs can be identified on the basis of amino acid sequence; thus, an analysis of the sequence of proteins may provide insight into both the enzymatic potential of the molecule and/or

molecules with which the protein may associate. One useful database is Pfam (protein families), which is a large collection of multiple sequence alignments and hidden Markov models covering many common protein domains. Versions are available via the internet from Washington University in St. Louis, the Sanger Center in England, and the Karolinska Institute in Sweden. See, e.g., Bateman, et al. (2000) Nuc. Acids Res. 28:263-266; Sonnhammer, et al. (1997) Proteins 28:405-420 ; Bateman, et al. (1999) Nuc. Acids Res. 27:260-262; and Sonnhammer, et al. (1998) Nuc. Acids Res. 26:320-322.

In another embodiment, the cancer sequences are transmembrane proteins. Transmembrane proteins are molecules that span a phospholipid bilayer of a cell. They may have an intracellular domain, an extracellular domain, or both. The intracellular domains of such proteins may have a number of functions including those already described for intracellular proteins. For example, the intracellular domain may have enzymatic activity and/or may serve as a binding site for additional proteins. Frequently the intracellular domain of transmembrane proteins serves both roles. For example certain receptor tyrosine kinases have both protein kinase activity and SH2 domains. In addition, autophosphorylation of tyrosines on the receptor molecule itself, creates binding sites for additional SH2 domain containing proteins.

Transmembrane proteins may contain from one to many transmembrane domains. For example, receptor tyrosine kinases, certain cytokine receptors, receptor guanylyl cyclases and receptor serine/threonine protein kinases contain a single transmembrane domain. However, various other proteins including channels and adenylyl cyclases contain numerous transmembrane domains. Many important cell surface receptors such as G protein coupled receptors (GPCRs) are classified as "seven transmembrane domain" proteins, as they contain 7 membrane spanning regions. Characteristics of transmembrane domains include approximately 17 consecutive hydrophobic amino acids that may be followed by charged amino acids. Therefore, upon analysis of the amino acid sequence of a particular protein, the localization and number of transmembrane domains within the protein may be predicted (see, e.g., PSORT web site http://psort.nibb.ac.jp/). Important transmembrane protein receptors include, but are not limited to the insulin receptor, insulin-like growth factor receptor, human growth hormone receptor, glucose transporters, transferrin receptor, epidermal growth factor receptor, low density lipoprotein receptor, epidermal growth factor receptor, leptin receptor, and interleukin receptors, e.g., IL-1 receptor, IL-2 receptor, etc.

The extracellular domains of transmembrane proteins are diverse; however, conserved motifs are found repeatedly among various extracellular domains. Conserved structure and/or functions have been ascribed to different extracellular motifs. Many extracellular domains are involved in binding to other molecules. In one aspect, extracellular domains are found on receptors. Factors that bind the receptor domain include circulating ligands, which may be peptides, proteins, or small molecules such as adenosine and the like. For example, growth factors such as EGF, FGF, and PDGF are circulating growth factors that bind to their cognate receptors to initiate a variety of cellular responses. Other factors include cytokines, mitogenic factors, neurotrophic factors, and the like. Extracellular domains also bind to cell-associated molecules. In this respect, they may mediate cell-cell interactions. Cell-associated ligands can be tethered to the cell, e.g., via a glycosylphosphatidylinositol (GPI) anchor, or may themselves be transmembrane proteins. Extracellular domains may also associate with the extracellular matrix and contribute to the maintenance of the cell structure.

Cancer proteins that are transmembrane are particularly preferred in the present invention as they are readily accessible targets for immunotherapeutics, as are described herein. In addition, as outlined below, transmembrane proteins can be also useful in imaging modalities. Antibodies may be used to label such readily accessible proteins in situ. Alternatively, antibodies can also label intracellular proteins, in which case samples are typically permeablized to provide access to intracellular proteins. In addition, some membrane proteins can be processed to release a soluble protein, or to expose a residual fragment. Released soluble proteins may be useful diagnostic markers, processed residual protein fragments may be useful lung markers of disease.

It will also be appreciated that a transmembrane protein can be made soluble by removing transmembrane sequences, e.g., through recombinant methods. Furthermore, transmembrane proteins that have been made soluble can be made to be secreted through recombinant means by adding an appropriate signal sequence.

In another embodiment, the cancer proteins are secreted proteins; the secretion of which can be either constitutive or regulated. These proteins may have a signal peptide or signal sequence that targets the molecule to the secretory pathway. Secreted proteins are involved in numerous physiological events; e.g., if circulating, they often serve to transmit signals to various other cell types. The secreted protein may function in an autocrine manner (acting on the cell that secreted the factor), a paracrine manner (acting on cells in

close proximity to the cell that secreted the factor), an endocrine manner (acting on cells at a distance, e.g, secretion into the blood stream), or exocrine (secretion, e.g., through a duct or to adjacent epithelial surface as sweat glands, sebaceous glands, pancreatic ducts, lacrimal glands, mammary glands, wax producing glands of the ear, etc.). Thus secreted molecules often find use in modulating or altering numerous aspects of physiology. Cancer proteins that are secreted proteins are particularly preferred in the present invention as they serve as good targets for diagnostic markers, e.g., for blood, plasma, serum, or stool tests. Those which are enzymes may be antibody or small molecule targets. Others may be useful as vaccine targets, e.g., via CTL mechanisms.

Use of cancer nucleic acids

As described above, cancer sequence is initially identified by substantial nucleic acid and/or amino acid sequence homology or linkage to the cancer sequences outlined herein. Such homology can be based upon the overall nucleic acid or amino acid sequence, and is generally determined as outlined below, using either homology programs or hybridization conditions. Typically, linked sequences on a mRNA are found on the same molecule.

As detailed elsewhere, percent identity can be determined using an algorithm such as BLAST. A preferred method utilizes the BLASTN module of WU-BLAST-2 set to the default parameters, with overlap span and overlap fraction set to 1 and 0.125, respectively. Alignment may include the introduction of gaps in the sequences to be aligned. In addition, for sequences which contain either more or fewer nucleotides than those of the nucleic acids described, the percentage of homology may be determined based on the number of homologous nucleosides in relation to the total number of nucleosides. Thus, e.g., homology of sequences shorter than those of the sequences identified will be determined using the number of nucleosides in the shorter sequence.

In one embodiment, the nucleic acid homology is determined through hybridization studies. Thus, e.g., nucleic acids which hybridize under high stringency to a described nucleic acid, or its complement, or is also found on naturally occurring mRNAs is considered a cancer sequence. In another embodiment, less stringent hybridization conditions are used; e.g., moderate or low stringency conditions may be used; see Ausubel, supra, and Tijssen, supra.

The cancer nucleic acid sequences of the invention, e.g., the sequences in Table 3, can be fragments of larger genes, e.g., they are nucleic acid segments. "Genes" in this

context includes coding regions, non-coding regions, and mixtures of coding and non-coding regions. Accordingly, using the sequences provided herein, extended sequences, in either direction, of the cancer genes can be obtained, using techniques well known for cloning either longer sequences or the full length sequences; see Ausubel, et al., supra. Much can be done by informatics and many sequences can be clustered to include multiple sequences corresponding to a single gene, e.g., systems such as UniGene (see, UniGene database at the NCBI web-site).

Once a cancer nucleic acid is identified, it can be cloned and, if necessary, its constituent parts recombined to form the entire cancer nucleic acid coding regions or the entire mRNA sequence. Once isolated from its natural source, e.g., contained within a plasmid or other vector or excised therefrom as a linear nucleic acid segment, the recombinant cancer nucleic acid can be further used as a probe to identify and isolate other cancer nucleic acids, e.g., extended coding regions. It can also be used as a "precursor" nucleic acid to make modified or variant cancer nucleic acids and proteins.

The cancer nucleic acids of the present invention are used in several ways. In one embodiment, nucleic acid probes to the cancer nucleic acids are made and attached to biochips to be used in screening and diagnostic methods, as outlined below, or for administration, e.g., for gene therapy, vaccine, RNAi, and/or antisense applications. Alternatively, cancer nucleic acids that include coding regions of cancer proteins can be put into expression vectors for the expression of cancer proteins, again for screening purposes or for administration to a patient.

In a preferred embodiment, nucleic acid probes to cancer nucleic acids (both the nucleic acid sequences outlined in the figures and/or the complements thereof) are made. The nucleic acid probes attached to the biochip are designed to be substantially complementary to the cancer nucleic acids, e.g., the target sequence (either the target sequence of the sample or to other probe sequences, e.g., in sandwich assays), such that hybridization of the target sequence and the probes of the present invention occurs. As outlined below, this complementarity need not be perfect; there may be any number of base pair mismatches which will interfere with hybridization between the target sequence and the single stranded nucleic acids of the present invention. However, if the number of mutations is so great that no hybridization can occur under even the least stringent of hybridization conditions, the sequence is not a complementary target sequence. Thus, by "substantially complementary" herein is meant that the probes are sufficiently complementary to the target

sequences to hybridize under normal reaction conditions, particularly high stringency conditions, as outlined herein.

A nucleic acid probe is generally single stranded but can be partially single and partially double stranded. The strandedness of the probe is dictated by the structure, composition, and properties of the target sequence. In general, the nucleic acid probes range from about 8-100 bases long, with from about 10-80 bases being preferred, and from about 30-50 bases being particularly preferred. That is, generally whole genes are not used. In some embodiments, much longer nucleic acids can be used, up to hundreds of bases.

In a preferred embodiment, more than one probe per sequence is used, with either overlapping probes or probes to different sections of the target being used. That is, two, three, four or more probes, with three being preferred, are used to build in a redundancy for a particular target. The probes can be overlapping (e.g., have some sequence in common), or separate. In some cases, PCR primers may be used to amplify signal for higher sensitivity.

Nucleic acids can be attached or immobilized to a solid support in a wide variety of ways. By "immobilized" and grammatical equivalents herein is meant the association or binding between the nucleic acid probe and the solid support is sufficient to be stable under the conditions of binding, washing, analysis, and removal as outlined. The binding can typically be covalent or non-covalent. By "non-covalent binding" and grammatical equivalents herein is meant one or more of electrostatic, hydrophilic, and hydrophobic interactions. Included in non-covalent binding is the covalent attachment of a molecule, e.g., streptavidin to the support and the non-covalent binding of the biotinylated probe to the streptavidin. By "covalent binding" and grammatical equivalents herein is meant that the two moieties, the solid support and the probe, are attached by at least one bond, including sigma bonds, pi bonds, and coordination bonds. Covalent bonds can be formed directly between the probe and the solid support or can be formed by a cross linker or by inclusion of a specific reactive group on either the solid support or the probe or both molecules. Immobilization may also involve a combination of covalent and non-covalent interactions.

In general, the probes are attached to the biochip in a wide variety of ways. As described herein, the nucleic acids can either be synthesized first, with subsequent attachment to the biochip, or can be directly synthesized on the biochip.

The biochip comprises a suitable solid substrate. By "substrate" or "solid support" or other grammatical equivalents herein is meant a material that can be modified for the

attachment or association of the nucleic acid probes and is amenable to at least one detection method. Often, the substrate may contain discrete individual sites appropriate for individual partitioning and identification. The number of possible substrates is very large, and include, but are not limited to, glass and modified or functionalized glass, plastics (including acrylics, polystyrene and copolymers of styrene and other materials, polypropylene, polyethylene, polybutylene, polyurethanes, TeflonJ, etc.), polysaccharides, nylon or nitrocellulose, resins, silica or silica-based materials including silicon and modified silicon, carbon, metals, inorganic glasses, plastics, etc. In general, the substrates allow optical detection and do not appreciably fluoresce. See WO 0055627.

Generally the substrate is planar, although other configurations of substrates may be used as well. For example, the probes may be placed on the inside surface of a tube for flow-through sample analysis to minimize sample volume. Similarly, the substrate may be flexible, such as a flexible foam, including closed cell foams made of particular plastics.

In a preferred embodiment, the surface of the biochip and the probe may be derivatized with chemical functional groups for subsequent attachment of the two. Thus, e.g., the biochip is derivatized with a chemical functional group including, but not limited to, amino groups, carboxy groups, oxo groups, and thiol groups, with amino groups being particularly preferred. Using these functional groups, the probes can be attached using functional groups on the probes. For example, nucleic acids containing amino groups can be attached to surfaces comprising amino groups, e.g., using linkers; e.g., homo-or hetero-bifunctional linkers as are well known (see 1994 Pierce Chemical Company catalog, technical section on cross-linkers, pages 155-200). In addition, in some cases, additional linkers, such as alkyl groups (including substituted and heteroalkyl groups) may be used.

In this embodiment, oligonucleotides are synthesized, and then attached to the surface of the solid support. Either the 5' or 3' terminus may be attached to the solid support, or attachment may be via linkage to an internal nucleoside. In another embodiment, the immobilization to the solid support may be very strong, yet non-covalent. For example, biotinylated oligonucleotides can be made, which bind to surfaces covalently coated with streptavidin, resulting in attachment.

Alternatively, the oligonucleotides may be synthesized on the surface. For example, photoactivation techniques utilizing photopolymerization compounds and techniques are used. In a preferred embodiment, the nucleic acids can be synthesized in situ, using known photolithographic techniques, such as those described in WO 95/25116; WO 95/35505; U.S.

Patent Nos. 5,700,637 and 5,445,934; and references cited within, all of which are expressly incorporated by reference; these methods of attachment form the basis of the Affymetrix GeneChip™ technology.

Often, amplification-based assays are performed to measure the expression level of cancer-associated sequences. These assays are typically performed in conjunction with reverse transcription. In such assays, a cancer-associated nucleic acid sequence acts as a template in an amplification reaction (e.g., Polymerase Chain Reaction, or PCR). In a quantitative amplification, the amount of amplification product will be proportional to the amount of template in the original sample. Comparison to appropriate controls provides a measure of the amount of cancer-associated RNA. Methods of quantitative amplification are well known. Detailed protocols for quantitative PCR are provided, e.g., in Innis, et al. (1990) PCR Protocols: A Guide to Methods and Applications Academic Press.

In some embodiments, a TaqMan based assay is used to measure expression. TaqMan based assays use a fluorogenic oligonucleotide probe that contains a 5' fluorescent dye and a 3' quenching agent. The probe hybridizes to a PCR product, but cannot itself be extended due to a blocking agent at the 3' end. When the PCR product is amplified in subsequent cycles, the 5' nuclease activity of the polymerase, e.g., AmpliTaq, results in the cleavage of the TaqMan probe. This cleavage separates the 5' fluorescent dye and the 3' quenching agent, thereby resulting in an increase in fluorescence as a function of amplification (see, e.g., literature provided by Perkin-Elmer at their public web site).

Other suitable amplification methods include, but are not limited to, ligase chain reaction (LCR) (see Wu and Wallace (1989) Genomics 4:560-569, Landegren, et al. (1988) Science 241:1077-1080, and Barringer, et al. (1990) Gene 89:117-122), transcription amplification (Kwoh, et al. (1989) Proc. Natl. Acad. Sci. USA 86:1173-1177), self-sustained sequence replication (Guatelli, et al. (1990) Proc. Natl. Acad. Sci. USA 87:1874-1878), dot PCR, linker adapter PCR, etc.

Expression of cancer proteins from nucleic acids

In a preferred embodiment, cancer nucleic acids, e.g., encoding cancer proteins, are used to make a variety of expression vectors to express cancer proteins which can then be used in screening assays, as described below. Expression vectors and recombinant DNA technology are well known (see, e.g., Ausubel, supra, and Fernandez and Hoeffler (eds. 1999) Gene Expression Systems Academic Press) to express proteins. The expression vectors may be either self-replicating extrachromosomal vectors or vectors which integrate

into a host genome. Generally, these expression vectors include transcriptional and translational regulatory nucleic acid operably linked to the nucleic acid encoding the cancer protein. The term "control sequences" refers to DNA sequences used for the expression of an operably linked coding sequence in a particular host organism. Control sequences that are suitable for prokaryotes, e.g., include a promoter, optionally an operator sequence, and a ribosome binding site. Eukaryotic cells are known to utilize promoters, polyadenylation signals, and enhancers.

Nucleic acid is "operably linked" when it is placed into a functional relationship with another nucleic acid sequence. For example, DNA for a presequence or secretory leader is operably linked to DNA for a polypeptide if it is expressed as a preprotein that participates in the secretion of the polypeptide; a promoter or enhancer is operably linked to a coding sequence if it affects the transcription of the sequence; or a ribosome binding site is operably linked to a coding sequence if it is positioned so as to facilitate translation. Generally, "operably linked" means that the DNA sequences being linked are contiguous, and, in the case of a secretory leader, contiguous and in reading phase. However, enhancers do not have to be contiguous. Linking is typically accomplished by ligation at convenient restriction sites. If such sites do not exist, synthetic oligonucleotide adaptors or linkers are used in accordance with conventional practice. Transcriptional and translational regulatory nucleic acid will generally be appropriate to the host cell used to express the cancer protein. Numerous types of appropriate expression vectors and suitable regulatory sequences are known for a variety of host cells.

In general, transcriptional and translational regulatory sequences may include, but are not limited to, promoter sequences, ribosomal binding sites, transcriptional start and stop sequences, translational start and stop sequences, and enhancer or activator sequences. In a preferred embodiment, the regulatory sequences include a promoter and transcriptional start and stop sequences.

Promoter sequences may be either constitutive or inducible promoters. The promoters may be either naturally occurring promoters or hybrid promoters. Hybrid promoters, which combine elements of more than one promoter, are also known, and are useful in the present invention.

An expression vector may comprise additional elements. For example, the expression vector may have two replication systems, thus allowing it to be maintained in two organisms, e.g., in mammalian or insect cells for expression and in a prokaryotic host

for cloning and amplification. Furthermore, for integrating expression vectors, the expression vector often contains at least one sequence homologous to the host cell genome, and preferably two homologous sequences which flank the expression construct. The integrating vector may be directed to a specific locus in the host cell by selecting the appropriate homologous sequence for inclusion in the vector. Constructs for integrating vectors are available. See, e.g., Fernandez and Hoeffler, supra; and Kitamura, et al. (1995) Proc. Nat'l Acad. Sci. USA 92:9146-9150.

In addition, in a preferred embodiment, the expression vector contains a selectable marker gene to allow the selection of transformed host cells. Selection genes are well known and will vary with the host cell used.

The cancer proteins of the present invention are usually produced by culturing a host cell transformed with an expression vector containing nucleic acid encoding a cancer protein, under the appropriate conditions to induce or cause expression of the cancer protein. Conditions appropriate for cancer protein expression will vary with the choice of the expression vector and the host cell, and will be easily ascertained through routine experimentation or optimization. For example, the use of constitutive promoters in the expression vector will require optimizing the growth and proliferation of the host cell, while the use of an inducible promoter requires the appropriate growth conditions for induction. In addition, in some embodiments, the timing of the harvest is important. For example, the baculoviral systems used in insect cell expression are lytic viruses, and thus harvest time selection can be crucial for product yield.

Appropriate host cells include yeast, bacteria, archaebacteria, fungi, and insect and animal cells, including mammalian cells. Of particular interest are Saccharomyces cerevisiae and other yeasts, E. coli, Bacillus subtilis, Sf9 cells, C129 cells, 293 cells, Neurospora, BHK, CHO, COS, HeLa cells, HUVEC (human umbilical vein endothelial cells), THP1 cells (a macrophage cell line), and various other human cells and cell lines.

In a preferred embodiment, the cancer proteins are expressed in mammalian cells. Mammalian expression systems may be used, and include retroviral and adenoviral systems. One expression vector system is a retroviral vector system such as is generally described in PCT/US97/01019 and PCT/US97/01048. Of particular use as mammalian promoters are the promoters from mammalian viral genes, since the viral genes are often highly expressed and have a broad host range. Examples include the SV40 early promoter, mouse mammary tumor virus LTR promoter, adenovirus major late promoter, herpes simplex virus promoter,

and the CMV promoter (see, e.g., Fernandez and Hoeffler, supra). Typically, transcription termination and polyadenylation sequences recognized by mammalian cells are regulatory regions located 3' to the translation stop codon and thus, together with the promoter elements, flank the coding sequence. Examples of transcription terminator and polyadenlyation signals include those derived from SV40.

Methods of introducing exogenous nucleic acid into mammalian hosts, as well as other hosts, are available, and will vary with the host cell used. Techniques include dextran-mediated transfection, calcium phosphate precipitation, polybrene mediated transfection, protoplast fusion, electroporation, viral infection, encapsulation of the polynucleotide(s) in liposomes, and direct microinjection of the DNA into nuclei.

In a preferred embodiment, cancer proteins are expressed in bacterial systems. Promoters from bacteriophage may also be used. In addition, synthetic promoters and hybrid promoters are also useful; e.g., the tac promoter is a hybrid of the trp and lac promoter sequences. Furthermore, a bacterial promoter can include naturally occurring promoters of non-bacterial origin that have the ability to bind bacterial RNA polymerase and initiate transcription. In addition to a functioning promoter sequence, an efficient ribosome binding site is desirable. The expression vector may also include a signal peptide sequence that provides for secretion of the cancer protein in bacteria. The protein is either secreted into the growth media (gram-positive bacteria) or into the periplasmic space, located between the inner and outer membrane of the cell (gram-negative bacteria). The bacterial expression vector may also include a selectable marker gene to allow for the selection of bacterial strains that have been transformed. Suitable selection genes include genes which render the bacteria resistant to drugs such as ampicillin, chloramphenicol, erythromycin, kanamycin, neomycin, and tetracycline. Selectable markers also include biosynthetic genes, such as those in the histidine, tryptophan, and leucine biosynthetic pathways. These components are assembled into expression vectors. Expression vectors for bacteria are well known, and include vectors for Bacillus subtilis, E. coli, Streptococcus cremoris, and Streptococcus lividans, among others (e.g., Fernandez and Hoeffler, supra). The bacterial expression vectors are transformed into bacterial host cells using techniques such as calcium chloride treatment, electroporation, and others.

In one embodiment, cancer proteins are produced in insect cells using, e.g., expression vectors for the transformation of insect cells, and in particular, baculovirus-based expression vectors.

In a preferred embodiment, a cancer protein is produced in yeast cells. Yeast expression systems are well known, and include expression vectors for Saccharomyces cerevisiae, Candida albicans and C. maltosa, Hansenula polymorpha, Kluyveromyces fragilis and K. lactis, Pichia guillerimondii and P. pastoris, Schizosaccharomyces pombe, and Yarrowia lipolytica.

The cancer protein may also be made as a fusion protein, using available techniques. Thus, e.g., for the creation of monoclonal antibodies, if the desired epitope is small, the cancer protein may be fused to a carrier protein to form an immunogen. Alternatively, the cancer protein may be made as a fusion protein to increase expression, or for other reasons. For example, when the cancer protein is a cancer peptide, the nucleic acid encoding the peptide may be linked to other nucleic acid for expression purposes. Fusion with detection epitope tags can be made, e.g., with FLAG, His6, myc, HA, etc.

In a preferred embodiment, the cancer protein is purified or isolated after expression. Cancer proteins may be isolated or purified in a variety of ways depending on what other components are present in the sample and the requirements for purified product, e.g., natural conformation or denatured. Standard purification methods include ammonium sulfate precipitations, electrophoretic, molecular, immunological, and chromatographic techniques, including ion exchange, hydrophobic, affinity, and reverse-phase HPLC chromatography, and chromatofocusing. For example, the cancer protein may be purified using a standard anti-cancer protein antibody column. Ultrafiltration and diafiltration techniques, in conjunction with protein concentration, are also useful. See, e.g., Walsh (2002) Proteins: Biochemistry and Biotechnology Wiley; Hardin, et al. (eds. 2001) Cloning, Gene Expression and Protein Purification Oxford Univ. Press; Wilson, et al. (eds. 2000) Encyclopedia of Separation Science Academic Press; and Scopes (1993) Protein Purification Springer-Verlag. The degree of purification necessary will vary depending on the use of the cancer protein. In some instances no purification will be necessary.

Once expressed and purified if necessary, the cancer proteins and nucleic acids are useful in a number of applications. They may be used as immunoselection reagents, as vaccine reagents, as screening agents, therapeutic entities, for production of antibodies, as transcription or translation inhibitors, etc.

Variants of cancer proteins

Also included within one embodiment of cancer proteins are amino acid variants of the naturally occurring sequences, as determined herein. Preferably, the variants are

preferably greater than about 75% homologous to the wild-type sequence, more preferably greater than about 80%, even more preferably greater than about 85%, and most preferably greater than 90%. In some embodiments the homology will be as high as about 93-95% or 98%. As for nucleic acids, homology in this context means sequence similarity or identity, with identity being preferred. This homology will be determined using standard techniques, as are outlined above for nucleic acid homologies.

Cancer proteins of the present invention may be shorter or longer than the wild type amino acid sequences. Thus, in a preferred embodiment, included within the definition of cancer proteins are portions or fragments of the wild type sequences herein. In addition, as outlined above, the cancer nucleic acids of the invention may be used to obtain additional coding regions, and thus additional protein sequence.

In one embodiment, the cancer proteins are derivative or variant cancer proteins as compared to the wild-type sequence. That is, as outlined more fully below, the derivative cancer peptide will often contain at least one amino acid substitution, deletion, or insertion, with amino acid substitutions being particularly preferred. The amino acid substitution, insertion, or deletion may occur at many residue positions within the cancer peptide.

Also included within one embodiment of cancer proteins of the present invention are amino acid sequence variants. These variants typically fall into one or more of three classes: substitutional, insertional, or deletional variants. These variants ordinarily are prepared by site specific mutagenesis of nucleotides in the DNA encoding the cancer protein, using cassette or PCR mutagenesis or other techniques, to produce DNA encoding the variant, and thereafter expressing the DNA in recombinant cell culture as outlined above. However, variant cancer protein fragments having up to about 100-150 residues may be prepared by in vitro synthesis using established techniques. Amino acid sequence variants are characterized by the predetermined nature of the variation, a feature that sets them apart from naturally occurring allelic or interspecies variation of the cancer protein amino acid sequence. The variants typically exhibit a similar qualitative biological activity as a naturally occurring analogue, although variants can also be selected which have modified characteristics.

While the site or region for introducing an amino acid sequence variation is often predetermined, the mutation per se need not be predetermined. For example, in order to optimize the performance of a mutation at a given site, random mutagenesis may be conducted at the target codon or region and the expressed cancer variants screened for the

optimal combination of desired activity. Techniques for making substitution mutations at predetermined sites in DNA having a known sequence are well known, e.g., M13 primer mutagenesis and PCR mutagenesis. Screening of mutants is often done using assays of cancer protein activities.

Amino acid substitutions are typically of single residues; insertions usually will be on the order of from about 1-20 amino acids, although considerably larger insertions may be tolerated. Deletions generally range from about 1-20 residues, although in some cases deletions may be much larger.

Substitutions, deletions, insertions, or combination thereof may be used to arrive at a final derivative. Generally these changes are done on a few amino acids to minimize the alteration of the molecule. However, larger changes may be tolerated in certain circumstances. When small alterations in the characteristics of the cancer protein are desired, substitutions are generally made in accordance with the amino acid substitution relationships described.

The variants typically exhibit essentially the same qualitative biological activity and will elicit the same immune response as a naturally-occurring analog, although variants also are selected to modify the characteristics of cancer proteins as needed. Alternatively, the variant may be designed such that a biological activity of the cancer protein is altered. For example, glycosylation sites may be added, altered, or removed.

Substantial changes in function or immunological identity are sometimes made by selecting substitutions that are less conservative than those described above. For example, substitutions may be made which more significantly affect: the structure of the polypeptide backbone in the area of the alteration, for example the alpha-helical or beta-sheet structure; the charge or hydrophobicity of the molecule at the target site; or the bulk of the side chain. Substitutions which generally are expected to produce the greatest changes in the polypeptide's properties are those in which (a) a hydrophilic residue, e.g., serine or threone is substituted for (or by) a hydrophobic residue, e.g., leucine, isoleucine, phenylalanine, valine, or alanine; (b) a cysteine or proline is substituted for (or by) another residue; (c) a residue having an electropositive side chain, e.g., lysine, arginine, or histidine, is substituted for (or by) an electronegative residue, e.g., glutamic or aspartic acid; (d) a residue having a bulky side chain, e.g., phenylalanine, is substituted for (or by) one not having a side chain, e.g., glycine; or (e) a proline residue is incorporated or substituted, which changes the degree of rotational freedom of the peptidyl bond.

Variants typically exhibit a similar qualitative biological activity and will elicit the same immune response as the naturally-occurring analog, although variants also are selected to modify the characteristics of the skin cancer proteins as needed. Alternatively, the variant may be designed such that the biological activity of the cancer protein is altered. For example, glycosylation sites may be altered or removed.

Covalent modifications of cancer polypeptides are included within the scope of this invention. One type of covalent modification includes reacting targeted amino acid residues of a cancer polypeptide with an organic derivatizing agent that is capable of reacting with selected side chains or the N-or C-terminal residues of a cancer polypeptide. Derivatization with bifunctional agents is useful, for instance, for crosslinking cancer polypeptides to a water-insoluble support matrix or surface for use in a method for purifying anti-cancer polypeptide antibodies or screening assays, as is more fully described below. Commonly used crosslinking agents include, e.g., 1,1-bis(diazoacetyl)-2-phenylethane, glutaraldehyde, N-hydroxysuccinimide esters, e.g., esters with 4-azidosalicylic acid, homobifunctional imidoesters, including disuccinimidyl esters such as 3,3'-dithiobis(succinimidylpropionate), bifunctional maleimides such as bis-N-maleimido-1,8-octane and agents such as methyl-3-((p-azidophenyl)dithio)propioimidate.

Other modifications include deamidation of glutaminyl and asparaginyl residues to the corresponding glutamyl and aspartyl residues, respectively, hydroxylation of proline and lysine, phosphorylation of hydroxyl groups of serinyl, threonyl, or tyrosyl residues, methylation of the amino groups of the lysine, arginine, and histidine side chains (e.g., pp. 79-86, Creighton (1992) Proteins: Structure and Molecular Properties Freeman), acetylation of the N-terminal amine, and amidation of a C-terminal carboxyl group.

Another type of covalent modification of the cancer polypeptide included within the scope of this invention comprises altering the native glycosylation pattern of the polypeptide. "Altering the native glycosylation pattern" is intended for purposes herein to mean deleting one or more carbohydrate moieties found in native sequence cancer polypeptide, and/or adding one or more glycosylation sites that are not present in the native sequence cancer polypeptide. Glycosylation patterns can be altered in many ways. Different cell types to express cancer-associated sequences can result in different glycosylation patterns.

Addition of glycosylation sites to cancer polypeptides may also be accomplished by altering the amino acid sequence thereof. The alteration may be made, e.g., by the addition

of, or substitution by, one or more serine or threonine residues to the native sequence cancer polypeptide (for O-linked glycosylation sites). The cancer amino acid sequence may optionally be altered through changes at the DNA level, particularly by mutating the DNA encoding the cancer polypeptide at preselected bases such that codons are generated that will translate into the desired amino acids.

Another means of increasing the number of carbohydrate moieties on the cancer polypeptide is by chemical or enzymatic coupling of glycosides to the polypeptide. See, e.g., WO 87/05330; pp. 259-306 in Aplin and Wriston (1981) CRC Crit. Rev. Biochem.

Removal of carbohydrate moieties present on the cancer polypeptide may be accomplished chemically or enzymatically or by mutational substitution of codons encoding for amino acid residues that serve as targets for glycosylation. Chemical deglycosylation techniques are applicable. See, e.g., Sojar and Bahl (1987) Arch. Biochem. Biophys. 259:52-57 and Edge, et al. (1981) Anal. Biochem. 118:131-137. Enzymatic cleavage of carbohydrate moieties on polypeptides can be achieved by the use of a variety of endo-and exo-glycosidases. See, e.g., Thotakura, et al. (1987) Meth. Enzymol. 138:350-359.

Another type of covalent modification of cancer comprises linking the cancer polypeptide to one of a variety of nonproteinaceous polymers, e.g., polyethylene glycol, polypropylene glycol, or polyoxyalkylenes, in the manner set forth in U.S. Patent Nos. 4,640,835; 4,496,689; 4,301,144; 4,670,417; 4,791,192, or 4,179,337.

Cancer polypeptides of the present invention may also be modified in a way to form chimeric molecules comprising a cancer polypeptide fused to another heterologous polypeptide or amino acid sequence. In one embodiment, such a chimeric molecule comprises a fusion of a cancer polypeptide with a tag polypeptide which provides an epitope to which an anti-tag antibody can selectively bind. The epitope tag is generally placed at the amino-or carboxyl-terminus of the cancer polypeptide. The presence of such epitope-tagged forms of a cancer polypeptide can be detected using an antibody against the tag polypeptide. Also, provision of the epitope tag enables the cancer polypeptide to be readily purified by affinity purification using an anti-tag antibody or another type of affinity matrix that binds to the epitope tag. In an alternative embodiment, the chimeric molecule may comprise a fusion of a cancer polypeptide with an immunoglobulin or a particular region of an immunoglobulin. For a bivalent form of the chimeric molecule, such a fusion could be to the Fc region of an IgG molecule.

Various tag polypeptides and their respective antibodies are available. Examples include poly-histidine (poly-his) or poly-histidine-glycine (poly-his-gly) tags; HIS6 and metal chelation tags, the flu HA tag polypeptide and its antibody 12CA5 (Field, et al. (1988) Mol. Cell. Biol. 8:2159-2165); the c-myc tag and the 8F9, 3C7, 6E10, G4, B7, and 9E10 antibodies thereto (Evan, et al. (1985) Molecular and Cellular Biology 5:3610-3616); and the Herpes Simplex virus glycoprotein D (gD) tag and its antibody (Paborsky, et al. (1990) Protein Engineering 3(6):547-553). Other tag polypeptides include the Flag-peptide (Hopp, et al. (1988) BioTechnology 6:1204-1210); the KT3 epitope peptide (Martin, et al. (1992) Science 255:192-194); tubulin epitope peptide (Skinner, et al. (1991) J. Biol. Chem. 266:15163-15166); and the T7 gene 10 protein peptide tag (Lutz-Freyermuth, et al. (1990) Proc. Natl. Acad. Sci. USA 87:6393-6397).

Also included are other cancer proteins of the cancer family, and cancer proteins from other organisms, which are cloned and expressed as outlined below. Thus, probe or degenerate polymerase chain reaction (PCR) primer sequences may be used to find other related cancer proteins from humans or other organisms. Particularly useful probe and/or PCR primer sequences include the unique areas of the cancer nucleic acid sequence. Preferred PCR primers are from about 15-35 nucleotides in length, with from about 20-30 being preferred, and may contain inosine as needed. The conditions for PCR reaction have been well described (e.g., Innis, PCR Protocols, supra).

In addition, cancer proteins can be made that are longer than those encoded by the nucleic acids of Table 2 or the attached listing of SEQ ID NOs:1-58, e.g., by the elucidation of extended sequences, the addition of epitope or purification tags, the addition of other fusion sequences, etc.

Cancer proteins may also be identified as being encoded by cancer nucleic acids. Thus, cancer proteins are encoded by nucleic acids that will hybridize to the sequences of the sequence listings, or their complements, as outlined herein.

Antibodies to cancer proteins

In a preferred embodiment, when the cancer protein is to be used to generate antibodies, e.g., for immunotherapy or immunodiagnosis, the cancer protein should share at least one epitope or determinant with the full length protein. By "epitope" or "determinant" herein is typically meant a portion of a protein which will generate and/or bind an antibody or T-cell receptor in the context of MHC. Thus, in most instances, antibodies made to a smaller cancer protein will be able to bind to the full-length protein, particularly linear

epitopes. In a preferred embodiment, the epitope is unique; that is, antibodies generated to a unique epitope show little or no cross-reactivity. In a preferred embodiment, the epitope is selected from a protein sequence set out in the Table 2 or the attached listing of SEQ ID NOs:59-116.

Methods of preparing polyclonal antibodies exist (e.g., Coligan, supra; and Harlow and Lane, supra). Polyclonal antibodies can be raised in a mammal, e.g., by one or more injections of an immunizing agent and, if desired, an adjuvant. Typically, the immunizing agent and/or adjuvant will be injected in the mammal by multiple subcutaneous or intraperitoneal injections. The immunizing agent may include a protein encoded by a nucleic acid of Table 2 or SEQ ID NOs:1-58 or fragment thereof or a fusion protein thereof. It may be useful to conjugate the immunizing agent to a protein known to be immunogenic in the mammal being immunized. Examples of such immunogenic proteins include but are not limited to keyhole limpet hemocyanin, serum albumin, bovine thyroglobulin, and soybean trypsin inhibitor. Examples of adjuvants which may be employed include Freund's complete adjuvant and MPL-TDM adjuvant (monophosphoryl Lipid A, synthetic trehalose dicorynomycolate). Various immunization protocols may be used.

The antibodies may, alternatively, be monoclonal antibodies. Monoclonal antibodies may be prepared using hybridoma methods, such as those described by Kohler and Milstein (1975) Nature 256:495. In a hybridoma method, a mouse, hamster, or other appropriate host animal, is typically immunized with an immunizing agent to elicit lymphocytes that produce or are capable of producing antibodies that will specifically bind to the immunizing agent. Alternatively, the lymphocytes may be immunized in vitro. The immunizing agent will typically include a polypeptide encoded by a nucleic acid of Table 2 or the attached listing of SEQ ID NOs:1-58, or fragment thereof, or a fusion protein thereof. Generally, either peripheral blood lymphocytes ("PBLs") are used if cells of human origin are desired, or spleen cells or lymph node cells are used if non-human mammalian sources are desired. The lymphocytes are then fused with an immortalized cell line using a suitable fusing agent, such as polyethylene glycol, to form a hybridoma cell (e.g., pp. 59-103 in Goding (1986) Monoclonal Antibodies: Principles and Practice Academic Press). Immortalized cell lines are usually transformed mammalian cells, particularly myeloma cells of rodent, bovine, or human origin. Usually, rat or mouse myeloma cell lines are employed. The hybridoma cells may be cultured in a suitable culture medium that preferably contains one or more substances that inhibit the growth or survival of the

unfused, immortalized cells. For example, if the parental cells lack the enzyme hypoxanthine guanine phosphoribosyl transferase (HGPRT or HPRT), the culture medium for the hybridomas typically will include hypoxanthine, aminopterin, and thymidine ("HAT medium"), which substances prevent the growth of HGPRT-deficient cells.

In one embodiment, the antibodies are bispecific antibodies. Bispecific antibodies are monoclonal, preferably human or humanized, antibodies that have binding specificities for at least two different antigens or that have binding specificities for two epitopes on the same antigen. In one embodiment, one of the binding specificities is for a protein encoded by a nucleic acid of Table 2 or the attached listing of SEQ ID NOs:1-58, or a fragment thereof, the other one is for another antigen, and preferably for a cell-surface protein or receptor or receptor subunit, preferably one that is tumor specific. Alternatively, tetramer-type technology may create multivalent reagents.

In a preferred embodiment, the antibodies to cancer protein are capable of reducing or eliminating a biological function of a cancer protein, in a naked form or conjugated to an effector moiety, as is described below. That is, the addition of anti-cancer protein antibodies (either polyclonal or preferably monoclonal) to cancer tissue (or cells containing cancer) may reduce or eliminate the cancer. Generally, at least a 25% decrease in activity, growth, size, or the like is preferred, with at least about 50% being particularly preferred and about a 95-100% decrease being especially preferred.

In a preferred embodiment the antibodies to the cancer proteins are humanized antibodies (e.g., Xenerex Biosciences, Medarex, Inc., Abgenix, Inc., Protein Design Labs, Inc.) Humanized forms of non-human (e.g., murine) antibodies are chimeric molecules of immunoglobulins, immunoglobulin chains or fragments thereof (such as Fv, Fab, Fab', F(ab')2 or other antigen-binding subsequences of antibodies) which contain minimal sequence derived from non-human immunoglobulin. Humanized antibodies include human immunoglobulins (recipient antibody) in which residues from a complementary determining region (CDR) of the recipient are replaced by residues from a CDR of a non-human species (donor antibody) such as mouse, rat, or rabbit having the desired specificity, affinity, and capacity. In some instances, Fv framework residues of a human immunoglobulin are replaced by corresponding non-human residues. Humanized antibodies may also comprise residues which are found neither in the recipient antibody nor in the imported CDR or framework sequences. In general, a humanized antibody will comprise substantially all of at least one, and typically two, variable domains, in which all or substantially all of the CDR

regions correspond to those of a non-human immunoglobulin and all or substantially all of the framework (FR) regions are those of a human immunoglobulin consensus sequence. The humanized antibody optimally also will typically comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin (Jones, et al. (1986) Nature 321:522-525; Riechmann, et al. (1988) Nature 332:323-329; and Presta (1992) Curr. Op. Struct. Biol. 2:593-596). Humanization can be essentially performed following the method of Winter and co-workers (Jones, et al. (1986) Nature 321:522-525; Riechmann, et al. (1988) Nature 332:323-327; Verhoeyen, et al. (1988) Science 239:1534-1536), by substituting rodent CDRs or CDR sequences for corresponding sequences of a human antibody. Accordingly, such humanized antibodies are chimeric antibodies (U.S. Patent No. 4,816,567), wherein substantially less than an intact human variable domain has been substituted by corresponding sequence from a non-human species.

Human antibodies can also be produced using phage display libraries (Hoogenboom and Winter (1992) J. Mol. Biol. 227:381-388; Marks, et al. (1991) J. Mol. Biol. 222:581-597) or human monoclonal antibodies (e.g., p. 77, Cole, et al. in Reisfeld and Sell (1985) Monoclonal Antibodies and Cancer Therapy Liss; and Boerner, et al. (1991) J. Immunol. 147:86-95). Similarly, human antibodies can be made by introducing human immunoglobulin loci into transgenic animals, e.g., mice in which the endogenous immunoglobulin genes have been partially or completely inactivated. Upon challenge, human antibody production is observed, which closely resembles that seen in humans in nearly all respects, including gene rearrangement, assembly, and antibody repertoire. This approach is described, e.g., in U.S. Patent Nos. 5,545,807; 5,545,806; 5,569,825; 5,625,126; 5,633,425; 5,661,016, and in the following scientific publications: Marks, et al. (1992) Bio/Technology 10:779-783; Lonberg, et al. (1994) Nature 368:856-859; Morrison (1994) Nature 368:812-13; Fishwild, et al. (1996) Nature Biotechnology 14:845-851; Neuberger (1996) Nature Biotechnology 14:826; and Lonberg and Huszar (1995) Intern. Rev. Immunol. 13:65-93.

By immunotherapy is meant treatment of cancer with an antibody raised against cancer proteins. As used herein, immunotherapy can be passive or active. Passive immunotherapy as defined herein is the passive transfer of antibody to a recipient (patient). Active immunization is the induction of antibody and/or T-cell responses in a recipient (patient). Induction of an immune response is the result of providing the recipient with an antigen to which antibodies are raised. The antigen may be provided by injecting a

polypeptide against which antibodies are desired to be raised into a recipient, or contacting the recipient with a nucleic acid capable of expressing the antigen and under conditions for expression of the antigen, leading to an immune response.

In a preferred embodiment the cancer proteins against which antibodies are raised are secreted proteins as described above. Without being bound by theory, antibodies used for treatment may bind and prevent the secreted protein from binding to its receptor, thereby inactivating the secreted cancer protein, e.g., in autocrine signaling.

In another preferred embodiment, the cancer protein to which antibodies are raised is a transmembrane protein. Without being bound by theory, antibodies used for treatment may bind the extracellular domain of the cancer protein and prevent it from binding to other proteins, such as circulating ligands or cell-associated molecules. The antibody may cause down-regulation of the transmembrane cancer protein. The antibody may be a competitive, non-competitive or uncompetitive inhibitor of protein binding to the extracellular domain of the cancer protein. The antibody may also be an antagonist of the cancer protein. Further, the antibody may prevent activation of the transmembrane cancer protein, or may induce or suppress a particular cellular pathway. In one aspect, when the antibody prevents the binding of other molecules to the cancer protein, the antibody prevents growth of the cell. The antibody may also be used to target or sensitize the cell to cytotoxic agents, including, but not limited to TNF-α, TNF-β, IL-1, INF-γ, and IL-2, or chemotherapeutic agents including 5FU, vinblastine, actinomycin D, cisplatin, methotrexate, and the like. In some instances the antibody may belong to a sub-type that activates serum complement when complexed with the transmembrane protein thereby mediating cytotoxicity or antigen-dependent cytotoxicity (ADCC). Thus, cancer may be treated by administering to a patient antibodies directed against the transmembrane cancer protein. Antibody-labeling may activate a co-toxin, localize a toxin payload, target a drug loaded liposome, or otherwise provide means to locally ablate cells.

In another preferred embodiment, the antibody is conjugated to an effector moiety. The effector moiety can be various molecules, including labeling moieties such as radioactive labels or fluorescent labels, or can be a therapeutic moiety. In one aspect the therapeutic moiety is a small molecule that modulates the activity of a cancer protein. In another aspect the therapeutic moiety may modulate the activity of molecules associated with or in close proximity to a cancer protein. The therapeutic moiety may inhibit enzymatic or signaling activity such as protease or collagenase or protein kinase activity

associated with cancer, or be an attractant of other cells, such as NK cells. See, e.g., USSN 09/544,494.

In a preferred embodiment, the therapeutic moiety can also be a cytotoxic agent. In this method, targeting the cytotoxic agent to cancer tissue or cells results in a reduction in the number of afflicted cells, thereby reducing symptoms associated with cancer. Cytotoxic agents are numerous and varied and include, but are not limited to, cytotoxic drugs or toxins or active fragments of such toxins. Suitable toxins and their corresponding fragments include diphtheria A chain, exotoxin A chain, ricin A chain, abrin A chain, curcin, crotin, phenomycin, enomycin, saporin, auristatin, and the like. Cytotoxic agents also include radiochemicals made by conjugating radioisotopes to antibodies raised against cancer proteins, or binding of a radionuclide to a chelating agent that has been covalently attached to the antibody. Targeting the therapeutic moiety to transmembrane cancer proteins not only serves to increase the local concentration of therapeutic moiety in the cancer afflicted area, but also serves to reduce deleterious side effects that may be associated with the untargeted therapeutic moiety. Antibody fragments may be used to target toxin loaded liposomes.

In another preferred embodiment, the cancer protein against which the antibodies are raised is an intracellular protein. In this case, the antibody may be conjugated to a protein which facilitates entry into the cell. In one case, the antibody enters the cell by endocytosis. In another embodiment, a nucleic acid encoding the antibody is administered to the individual or cell. Moreover, wherein the cancer protein can be targeted within a cell, e.g., the nucleus, an antibody thereto may contain a signal for that target localization, e.g., a nuclear localization signal.

The cancer antibodies of the invention specifically bind to cancer proteins. By "specifically bind" herein is meant that the antibodies bind to the protein with a K_d of at least about 0.1 mM, more usually at least about 1 μM, preferably at least about 0.1 μM or better, and most preferably, 0.01 μM or better. Selectivity of binding to the specific target and not to related sequences is often also important.

Detection of cancer sequence for diagnostic and therapeutic applications

In one aspect, the RNA expression levels of genes are determined for different cellular states in the cancer phenotype. Expression levels of genes in normal tissue (e.g., not undergoing cancer) and in cancer tissue (and in some cases, for varying severities of cancer that relate to prognosis, as outlined below), or in non-malignant disease are evaluated

to provide expression profiles. A gene expression profile of a particular cell state or point of development is essentially a "fingerprint" of the state of the cell. While two states may have a particular gene similarly expressed, the evaluation of a number of genes simultaneously allows the generation of a gene expression profile that is reflective of the state of the cell. By comparing expression profiles of cells in different states, information regarding which genes are important (including both up- and down-regulation of genes) in each of these states is obtained. Then, diagnosis may be performed or confirmed to determine whether a tissue sample has the gene expression profile of normal or cancerous tissue. This will provide for molecular diagnosis of related conditions.

"Differential expression," or grammatical equivalents as used herein, refers to qualitative or quantitative differences in the temporal and/or cellular gene expression patterns within and among cells and tissue. Thus, a differentially expressed gene can qualitatively have its expression altered, including an activation or inactivation, in, e.g., normal versus cancer tissue. Genes may be turned on or turned off in a particular state, relative to another state thus permitting comparison of two or more states. A qualitatively regulated gene will exhibit an expression pattern within a state or cell type which is detectable by standard techniques. Some genes will be expressed in one state or cell type, but not in both. Alternatively, the difference in expression may be quantitative, e.g., in that expression is increased or decreased; e.g., gene expression is either upregulated, resulting in an increased amount of transcript, or downregulated, resulting in a decreased amount of transcript. The degree to which expression differs need only be large enough to quantify via standard characterization techniques as outlined below, such as by use of Affymetrix GeneChip™ expression arrays. See, Lockhart (1996) Nature Biotechnology 14:1675-1680. Other techniques include, but are not limited to, quantitative reverse transcriptase PCR, northern analysis, and RNase protection. As outlined above, preferably the change in expression (e.g., upregulation or downregulation) is at least about 50%, more preferably at least about 100%, more preferably at least about 150%, more preferably at least about 200%, with from 300 to at least 1000% being especially preferred.

Evaluation may be at the gene transcript or the protein level. The amount of gene expression may be monitored using nucleic acid probes to the RNA or DNA equivalent of the gene transcript, and the quantification of gene expression levels, or, alternatively, the final gene product itself (protein) can be monitored, e.g., with antibodies to the cancer protein and standard immunoassays (ELISAs, etc.) or other techniques, including mass

spectroscopy assays, 2D gel electrophoresis assays, etc. Proteins corresponding to cancer genes, e.g., those identified as being important in a cancer or disease phenotype, can be evaluated in a cancer diagnostic test. In a preferred embodiment, gene expression monitoring is performed simultaneously on a number of genes. Multiple protein expression monitoring can be performed as well.

In this embodiment, the cancer nucleic acid probes are attached to biochips as outlined herein for the detection and quantification of cancer sequences in a particular cell. The assays are further described below in the example. PCR techniques can be used to provide greater sensitivity.

In a preferred embodiment nucleic acids encoding the cancer protein are detected. Although DNA or RNA encoding the cancer protein may be detected, of particular interest are methods wherein an mRNA encoding a cancer protein is detected. Probes to detect mRNA can be a nucleotide/deoxynucleotide probe that is complementary to and hybridizes with the mRNA and includes, but is not limited to, oligonucleotides, cDNA, or RNA. Probes also should contain a detectable label, as defined herein. In one method the mRNA is detected after immobilizing the nucleic acid to be examined on a solid support such as nylon membranes and hybridizing the probe with the sample. Following washing to remove the non-specifically bound probe, the label is detected. In another method, detection of the mRNA is performed in situ. In this method permeabilized cells or tissue samples are contacted with a detectably labeled nucleic acid probe for sufficient time to allow the probe to hybridize with the target mRNA. Following washing to remove the non-specifically bound probe, the label is detected. For example a digoxygenin labeled riboprobe (RNA probe) that is complementary to the mRNA encoding a cancer protein is detected by binding the digoxygenin with an anti-digoxygenin secondary antibody and developed with nitro blue tetrazolium and 5-bromo-4-chloro-3-indoyl phosphate.

In a preferred embodiment, various proteins from the three classes of proteins as described herein (secreted, transmembrane, or intracellular proteins) are used in diagnostic assays. The cancer proteins, antibodies, nucleic acids, modified proteins, and cells containing cancer sequences are used in diagnostic assays. This can be performed on an individual gene or corresponding polypeptide level. In a preferred embodiment, the expression profiles are used, preferably in conjunction with high throughput screening techniques to allow monitoring for expression profile genes and/or corresponding polypeptides.

As described and defined herein, cancer proteins, including intracellular, transmembrane, or secreted proteins, find use as markers of cancer, e.g., for prognostic or diagnostic purposes. Detection of these proteins in putative cancer tissue allows for detection, prognosis, or diagnosis of cancer or similar disease, and for selection of therapeutic strategy. In one embodiment, antibodies are used to detect cancer proteins. A preferred method separates proteins from a sample by electrophoresis on a gel (typically a denaturing and reducing protein gel, but may be another type of gel, including isoelectric focusing gels and the like). Following separation of proteins, the cancer protein is detected, e.g., by immunoblotting with antibodies raised against the cancer protein.

In another preferred method, antibodies to the cancer protein find use in in situ imaging techniques, e.g., in histology. See, e.g., Asai, et al. (eds. 1993) Methods in Cell Biology: Antibodies in Cell Biology (vol. 37) Academic Press. In this method, cells are contacted with from one to many antibodies to the cancer protein(s). Following washing to remove non-specific antibody binding, the presence of the antibody or antibodies is detected. In one embodiment the antibody is detected by incubating with a secondary antibody that contains a detectable label. In another method the primary antibody to the cancer protein(s) contains a detectable label, e.g., an enzyme marker that can act on a substrate. In another preferred embodiment each one of multiple primary antibodies contains a distinct and detectable label. This method finds particular use in simultaneous screening for a plurality of cancer proteins. Many other histological imaging techniques are also provided by the invention.

In a preferred embodiment the label is detected in a fluorometer which has the ability to detect and distinguish emissions of different wavelengths. In addition, a fluorescence activated cell sorter (FACS) can be used in the method.

In another preferred embodiment, antibodies find use in diagnosing cancer from blood, serum, plasma, stool, and other samples. Such samples, therefore, are useful as samples to be probed or tested for the presence of cancer proteins. Antibodies can be used to detect a cancer protein by previously described immunoassay techniques including ELISA, immunoblotting (western blotting), immunoprecipitation, BIACORE technology and the like. Conversely, the presence of antibodies may indicate an immune response against an endogenous cancer protein.

In a preferred embodiment, in situ hybridization of labeled cancer nucleic acid probes to tissue arrays is done. For example, arrays of tissue samples, including cancer

tissue and/or normal tissue, are made. In situ hybridization (see, e.g., Ausubel, supra) is then performed. When comparing the fingerprints between an individual and a standard, a diagnosis, a prognosis, or a prediction may be based on the findings. It is further understood that the genes which indicate the diagnosis may differ from those which indicate the prognosis and molecular profiling of the condition of the cells may lead to distinctions between responsive or refractory conditions or may be predictive of outcomes.

In a preferred embodiment, the cancer proteins, antibodies, nucleic acids, modified proteins, and cells containing cancer sequences are used in prognosis assays. As above, gene expression profiles can be generated that correlate to cancer, clinical, pathological, or other information, in terms of long term prognosis. Again, this may be done on either a protein or gene level, with the use of genes being preferred. Single or multiple genes may be useful in various combinations. As above, cancer probes may be attached to biochips for the detection and quantification of cancer sequences in a tissue or patient. The assays proceed as outlined above for diagnosis. PCR method may provide more sensitive and accurate quantification.

Assays for therapeutic compounds

In a preferred embodiment, the proteins, nucleic acids, and antibodies as described herein are used in drug screening assays. The cancer proteins, antibodies, nucleic acids, modified proteins, and cells containing cancer sequences are used in drug screening assays or by evaluating the effect of drug candidates on a "gene expression profile" or expression profile of polypeptides. In a preferred embodiment, the expression profiles are used, preferably in conjunction with high throughput screening techniques, to allow monitoring for expression profile genes after treatment with a candidate agent (e.g., Zlokarnik, et al. (1998) Science 279:84-88; Heid (1996) Genome Res. 6:986-994.

In a preferred embodiment, the cancer proteins, antibodies, nucleic acids, modified proteins and cells containing the native or modified cancer proteins are used in screening assays. That is, the present invention provides novel methods for screening for compositions which modulate the cancer phenotype or an identified physiological function of a cancer protein. As above, this can be done on an individual gene level or by evaluating the effect of drug candidates on a "gene expression profile". In a preferred embodiment, the expression profiles are used, preferably in conjunction with high throughput screening techniques, to allow monitoring for expression profile genes after treatment with a candidate agent, see Zlokarnik, supra.

Having identified the differentially expressed genes herein, a variety of assays may be performed. In a preferred embodiment, assays may be run on an individual gene or protein level. That is, having identified a particular gene as up regulated in cancer, test compounds can be screened for the ability to modulate gene expression or for binding to the cancer protein. "Modulation" thus includes both an increase and a decrease in gene expression. The preferred amount of modulation will depend on the original change of the gene expression in normal versus tissue undergoing cancer, with changes of at least 10%, preferably 50%, more preferably 100-300%, and in some embodiments 300-1000% or greater. Thus, if a gene exhibits a 4-fold increase in cancer tissue compared to normal tissue, a decrease of about four-fold is often desired; similarly, a 10-fold decrease in cancer tissue compared to normal tissue often provides a target value of a 10-fold increase in expression to be induced by the test compound.

The amount of gene expression may be monitored using nucleic acid probes and the quantification of gene expression levels, or, alternatively, the gene product itself can be monitored, e.g., through the use of antibodies to the cancer protein and standard immunoassays. Proteomics and separation techniques may also allow quantification of expression.

In a preferred embodiment, gene expression or protein monitoring of a number of entities, e.g., an expression profile, is monitored simultaneously. Such profiles will typically involve a plurality of those entities described herein.

In this embodiment, the cancer nucleic acid probes are attached to biochips as outlined herein for the detection and quantification of cancer sequences in a particular cell. Alternatively, PCR may be used. Thus, a series, e.g., of microtiter plate, may be used with dispensed primers in desired wells. A PCR reaction can then be performed and analyzed for each well.

Modulators of cancer

Expression monitoring can be performed to identify compounds that modify the expression of one or more cancer-associated sequences, e.g., a polynucleotide sequence set out in Table 2 or SEQ ID NOs:1-58. Generally, in a preferred embodiment, a test modulator is added to the cells prior to analysis. Moreover, screens are also provided to identify agents that modulate cancer, modulate cancer proteins, bind to a cancer protein, or interfere with the binding of a cancer protein and an antibody or other binding partner.

The term "test compound" or "drug candidate" or "modulator" or grammatical equivalents as used herein describes a molecule, e.g., protein, oligopeptide, small organic molecule, polysaccharide, polynucleotide, etc., to be tested for the capacity to directly or indirectly alter the cancer phenotype or the expression of a cancer sequence, e.g., a nucleic acid or protein sequence. In preferred embodiments, modulators alter expression profiles, or expression profile nucleic acids or proteins provided herein. In one embodiment, the modulator suppresses a cancer phenotype, e.g., to a normal or non-malignant tissue fingerprint. In another embodiment, a modulator induced a cancer phenotype. Generally, a plurality of assay mixtures are run in parallel with different agent concentrations to obtain a differential response to the various concentrations. Typically, one of these concentrations serves as a negative control, e.g., at zero concentration or below the level of detection.

Drug candidates encompass numerous chemical classes, though typically they are organic molecules, preferably small organic compounds having a molecular weight of more than 100 and less than about 2,500 daltons. Preferred small molecules are less than 2000, or less than 1500, or less than 1000, or less than 500 D. Candidate agents comprise functional groups necessary for structural interaction with proteins, particularly hydrogen bonding, and typically include at least an amine, carbonyl, hydroxyl or carboxyl group, preferably at least two of the functional chemical groups. The candidate agents often comprise cyclical carbon or heterocyclic structures and/or aromatic or polyaromatic structures substituted with one or more of the above functional groups. Candidate agents are also found among biomolecules including peptides, saccharides, fatty acids, steroids, purines, pyrimidines, derivatives, structural analogs, or combinations thereof. Particularly preferred are peptides.

In one aspect, a modulator will neutralize the effect of a cancer protein. By "neutralize" is meant that activity of a protein is inhibited or blocked and the consequent effect on the cell.

In certain embodiments, combinatorial libraries of potential modulators will be screened for an ability to bind to a cancer polypeptide or to modulate activity. Conventionally, new chemical entities with useful properties are generated by identifying a chemical compound (called a "lead compound") with some desirable property or activity, e.g., inhibiting activity, creating variants of the lead compound, and evaluating the property and activity of those variant compounds. Often, high throughput screening (HTS) methods are employed for such an analysis. See, e.g., Janzen (2002) High Throughput Screening: Methods and Protocols Humana; Devlin (ed. 1997) High Throughput Screening: The

Discovery of Bioactive Substances Dekker; and Mei and Czarnik (eds. 2002) Integrated Drug Discovery Techniques Dekker.

In one preferred embodiment, high throughput screening methods involve providing a library containing a large number of potential therapeutic compounds (candidate compounds). Such "combinatorial chemical libraries" are then screened in one or more assays to identify those library members (particular chemical species or subclasses) that display a desired characteristic activity. The compounds thus identified can serve as conventional "lead compounds" or can themselves be used as potential or actual therapeutics.

A combinatorial chemical library is a collection of diverse chemical compounds generated by either chemical synthesis or biological synthesis by combining a number of chemical "building blocks" such as reagents. For example, a linear combinatorial chemical library, such as a polypeptide (e.g., mutein) library, is formed by combining a set of chemical building blocks called amino acids in every possible way for a given compound length (e.g., the number of amino acids in a polypeptide compound). Millions of chemical compounds can be synthesized through such combinatorial mixing of chemical building blocks (Gallop, et al. (1994) J. Med. Chem. 37:1233-1251).

Preparation and screening of combinatorial chemical libraries is well known. Such combinatorial chemical libraries include, but are not limited to, peptide libraries (see, e.g., U.S. Patent No. 5,010,175, Furka (1991) Pept. Prot. Res. 37:487-493, Houghton, et al. (1991) Nature 354:84-88), peptoids (PCT Publication No WO 91/19735), encoded peptides (PCT Publication WO 93/20242), random bio-oligomers (PCT Publication WO 92/00091), benzodiazepines (U.S. Pat. No. 5,288,514), diversomers such as hydantoins, benzodiazepines and dipeptides (Hobbs, et al. (1993) Proc. Natl. Acad. Sci. USA 90:6909-6913, vinylogous polypeptides (Hagihara, et al. (1992) J. Amer. Chem. Soc. 114:6568-570), nonpeptidal peptidomimetics with a Beta-D-Glucose scaffolding (Hirschmann, et al. (1992) J. Amer. Chem. Soc. 114:9217-9218), analogous organic syntheses of small compound libraries (Chen, et al. (1994) J. Amer. Chem. Soc. 116:2661-662), oligocarbamates (Cho, et al. (1993) Science 261:1303-1305), and/or peptidyl phosphonates (Campbell, et al. (1994) J. Org. Chem. 59:658). See, generally, Gordon, et al. (1994) J. Med. Chem. 37:1385-1401, nucleic acid libraries (see, e.g., Stratagene, Corp.), peptide nucleic acid libraries (see, e.g., U.S. Patent 5,539,083), antibody libraries (see, e.g., Vaughn, et al. (1996) Nature Biotechnology 14(3):309-314, and PCT/US96/10287), carbohydrate libraries (see, e.g.,

Liang, et al. (1996) Science 274:1520-1522, and U.S. Patent No. 5,593,853), and small organic molecule libraries (see, e.g., benzodiazepines, page 33 Baum (Jan 18, 1993) C&EN; isoprenoids, U.S. Patent No. 5,569,588; thiazolidinones and metathiazanones, U.S. Patent No. 5,549,974; pyrrolidines, U.S. Patent Nos. 5,525,735 and 5,519,134; morpholino compounds, U.S. Patent No. 5,506,337; benzodiazepines, U.S. Patent No. 5,288,514; and the like).

Devices for the preparation of combinatorial libraries are commercially available (see, e.g., 357 MPS, 390 MPS, Advanced Chem Tech, Louisville KY, Symphony, Rainin, Woburn, MA, 433A Applied Biosystems, Foster City, CA, 9050 Plus, Millipore, Bedford, MA).

A number of well known robotic systems have also been developed for solution phase chemistries. These systems include automated workstations like the automated synthesis apparatus developed by Takeda Chemical Industries, LTD. (Osaka, Japan) and many robotic systems utilizing robotic arms (Zymate II, Zymark Corporation, Hopkinton, Mass.; Orca, Hewlett-Packard, Palo Alto, Calif.), which mimic manual synthetic operations performed by a chemist. The above devices are suitable for use with the present invention. The nature and implementation of modifications to these devices (if any) so that they can operate as discussed herein will be apparent. In addition, numerous combinatorial libraries are themselves commercially available (see, e.g., ComGenex, Princeton, N.J., Asinex, Moscow, Ru, Tripos, Inc., St. Louis, MO, ChemStar, Ltd, Moscow, RU, 3D Pharmaceuticals, Exton, PA, Martek Biosciences, Columbia, MD, etc.).

The assays to identify modulators are amenable to high throughput screening. Preferred assays thus detect enhancement or inhibition of cancer gene transcription, inhibition, or enhancement of polypeptide expression, and inhibition or enhancement of polypeptide activity.

High throughput assays for the presence, absence, quantification, or other properties of particular nucleic acids or protein products are well known. Similarly, binding assays and reporter gene assays are similarly well known. Thus, e.g., U.S. Patent No. 5,559,410 discloses high throughput screening methods for proteins, U.S. Patent No. 5,585,639 discloses high throughput screening methods for nucleic acid binding (e.g., in arrays), while U.S. Patent Nos. 5,576,220 and 5,541,061 disclose high throughput methods of screening for ligand/antibody binding.

In addition, high throughput screening systems are commercially available (see, e.g., Zymark Corp., Hopkinton, MA; Air Technical Industries, Mentor, OH; Beckman Instruments, Inc. Fullerton, CA; Precision Systems, Inc., Natick, MA, etc.). These systems typically automate entire procedures, including sample and reagent pipetting, liquid dispensing, timed incubations, and final readings of the microplate in detector(s) appropriate for the assay. These configurable systems provide high throughput and rapid start up as well as a high degree of flexibility and customization. The manufacturers of such systems provide detailed protocols for various high throughput systems. Thus, e.g., Zymark Corp. provides technical bulletins describing screening systems for detecting the modulation of gene transcription, ligand binding, and the like.

In one embodiment, modulators are proteins, often naturally occurring proteins or fragments of naturally occurring proteins. Thus, e.g., cellular extracts containing proteins, or random or directed digests of proteinaceous cellular extracts, may be used. In this way libraries of proteins may be made for screening in the methods of the invention. Particularly preferred in this embodiment are libraries of bacterial, fungal, viral, and mammalian proteins, with the latter being preferred, and human proteins being especially preferred. Particularly useful test compound will be directed to the class of proteins to which the target belongs, e.g., substrates for enzymes or ligands and receptors.

In a preferred embodiment, modulators are peptides of from about 5-30 amino acids, with from about 5-20 amino acids being preferred, and from about 7-15 being particularly preferred. The peptides may be digests of naturally occurring proteins, random peptides, or "biased" random peptides. By "randomized" or grammatical equivalents herein is meant that each nucleic acid and peptide consists of essentially random nucleotides and amino acids, respectively. Since generally these random peptides (or nucleic acids, discussed below) are chemically synthesized, they may incorporate a nucleotide or amino acid at any position. The synthetic process can be designed to generate randomized proteins or nucleic acids, to allow the formation of all or most of the possible combinations over the length of the sequence, thus forming a library of randomized candidate bioactive proteinaceous agents.

In one embodiment, the library is fully randomized, with no sequence preferences or constants at any position. In a preferred embodiment, the library is biased. That is, some positions within the sequence are either held constant, or are selected from a limited number of possibilities. For example, in a preferred embodiment, the nucleotides or amino acid

residues are randomized within a defined class, e.g., of hydrophobic amino acids, hydrophilic residues, sterically biased (either small or large) residues, towards the creation of nucleic acid binding domains, the creation of cysteines, for cross-linking, prolines for SH-3 domains, serines, threonines, tyrosines, or histidines for phosphorylation sites, etc., or to purines, etc.

Modulators of cancer can also be nucleic acids, as defined above.

As described above generally for proteins, nucleic acid modulating agents may be naturally occurring nucleic acids, random nucleic acids, or "biased" random nucleic acids. For example, digests of prokaryotic or eukaryotic genomes may be used as is outlined above for proteins.

In a preferred embodiment, the candidate compounds are organic chemical moieties, a wide variety of which are available in the literature.

After the candidate agent has been added and the cells allowed to incubate for some period of time, the sample containing a target sequence to be analyzed is added to the biochip. If required, the target sequence is prepared using known techniques. For example, the sample may be treated to lyse the cells, using known lysis buffers, electroporation, etc., with purification and/or amplification such as PCR performed as appropriate. For example, an in vitro transcription with labels covalently attached to the nucleotides is performed. Generally, the nucleic acids are labeled with biotin-FITC or PE, or with cy3 or cy5.

In a preferred embodiment, the target sequence is labeled with, e.g., a fluorescent, a chemiluminescent, a chemical, or a radioactive signal, to provide a means of detecting the target sequence's specific binding to a probe. The label also can be an enzyme, such as, alkaline phosphatase or horseradish peroxidase, which when provided with an appropriate substrate produces a product that can be detected. Alternatively, the label can be a labeled compound or small molecule, such as an enzyme inhibitor, that binds but is not catalyzed or altered by the enzyme. The label also can be a moiety or compound, such as, an epitope tag or biotin which specifically binds to streptavidin. For the example of biotin, the streptavidin is labeled as described above, thereby, providing a detectable signal for the bound target sequence. Unbound labeled streptavidin is typically removed prior to analysis.

These assays can be direct hybridization assays or can comprise "sandwich assays", which include the use of multiple probes, as is generally outlined in U.S. Patent Nos. 5,681,702, 5,597,909, 5,545,730, 5,594,117, 5,591,584, 5,571,670, 5,580,731, 5,571,670, 5,591,584, 5,624,802, 5,635,352, 5,594,118, 5,359,100, 5,124,246, and 5,681,697, all of

which are hereby incorporated by reference. In this embodiment, in general, the target nucleic acid is prepared as outlined above, and then added to the biochip comprising a plurality of nucleic acid probes, under conditions that allow the formation of a hybridization complex.

A variety of hybridization conditions may be used in the present invention, including high, moderate, and low stringency conditions as outlined above. The assays are generally run under stringency conditions which allows formation of the label probe hybridization complex only in the presence of target. Stringency can be controlled by altering a step parameter that is a thermodynamic variable, including, but not limited to, temperature, formamide concentration, salt concentration, chaotropic salt concentration, pH, organic solvent concentration, etc.

These parameters may also be used to control non-specific binding, as is generally outlined in U.S. Patent No. 5,681,697. Thus it may be desirable to perform certain steps at higher stringency conditions to reduce non-specific binding.

The reactions outlined herein may be accomplished in a variety of ways. Components of the reaction may be added simultaneously, or sequentially, in different orders, with preferred embodiments outlined below. In addition, the reaction may include a variety of other reagents. These include salts, buffers, neutral proteins, e.g., albumin, detergents, etc. which may be used to facilitate optimal hybridization and detection, and/or reduce non-specific or background interactions. Reagents that otherwise improve the efficiency of the assay, such as protease inhibitors, nuclease inhibitors, anti-microbial agents, etc., may also be used as appropriate, depending on the sample preparation methods and purity of the target.

The assay data are analyzed to determine the expression levels, and changes in expression levels as between states of individual genes, forming a gene expression profile.

Screens are performed to identify modulators of the cancer phenotype. In one embodiment, screening is performed to identify modulators that can induce or suppress a particular expression profile, thus preferably generating the associated phenotype. In another embodiment, e.g., for diagnostic applications, having identified differentially expressed genes important in a particular state, screens can be performed to identify modulators that alter expression of individual genes. In an another embodiment, screening is performed to identify modulators that alter a biological function of the expression product of a differentially expressed gene. Again, having identified the importance of a gene in a

particular state, screens are performed to identify agents that bind and/or modulate the biological activity of the gene product.

In addition, screens can be done for genes that are induced in response to a candidate agent or treatment process. After identifying a modulator based upon its ability to suppress a cancer expression pattern leading to a normal expression pattern (or its converse), or to modulate a single cancer gene expression profile so as to mimic the expression of the gene from normal tissue, a screen as described above can be performed to identify genes that are specifically modulated in response to the agent. Comparing expression profiles between normal tissue and agent treated cancer tissue reveals genes that are not expressed in normal tissue or cancer tissue, but are expressed in agent treated tissue. These agent-specific sequences can be identified and used by methods described herein for cancer genes or proteins. In particular, these sequences and the proteins they encode find use in marking or identifying agent treated cells. In addition, antibodies can be raised against the agent induced proteins and used to target novel therapeutics, e.g., toxin loaded liposomes, to the treated cancer tissue sample.

Thus, in one embodiment, a test compound is administered to a population of cancer cells that have an associated cancer expression profile. By "administration" or "contacting" herein is meant that the candidate agent is added to the cells in such a manner as to allow the agent to act upon the cell, whether by uptake and intracellular action, or by action at the cell surface. In some embodiments, nucleic acid encoding a proteinaceous candidate agent (e.g., a peptide) may be put into a viral construct such as an adenoviral or retroviral construct, and added to the cell, such that expression of the peptide agent is accomplished, e.g., PCT US97/01019. Regulatable gene therapy systems can also be used.

Once a test compound has been administered to the cells, the cells can be washed if desired and are allowed to incubate under preferably physiological conditions for some period of time. The cells are then harvested and a new gene expression profile is generated, as outlined herein.

Thus, e.g., cancer or non-malignant tissue may be screened for agents that modulate, e.g., induce or suppress a cancer phenotype. A change in at least one gene, preferably many, of the expression profile indicates that the agent has an effect on cancer activity. By defining such a signature for the cancer phenotype, screens for new drugs that alter the phenotype can be devised. With this approach, the drug target need not be known and need

not be represented in the original expression screening platform, nor does the level of transcript for the target protein need to change.

In a preferred embodiment, as outlined above, screens may be done on individual genes and gene products (proteins). That is, having identified a particular differentially expressed gene as important in a particular state, screening of modulators of either the expression of the gene or the gene product itself can be done. The gene products of differentially expressed genes are sometimes referred to herein as "cancer proteins" or a "cancer modulatory protein". The cancer modulatory protein may be a fragment, or alternatively, be the full length protein to the fragment encoded by the nucleic acids of Table 2 or SEQ ID NOs:1-58. Preferably, the cancer modulatory protein is a fragment. In a preferred embodiment, the cancer amino acid sequence which is used to determine sequence identity or similarity is encoded by a nucleic acid of the Table 2 or SEQ ID NOs:1-58. In another embodiment, the sequences are naturally occurring allelic variants of a protein encoded by a nucleic acid of the Table 2 or SEQ ID NOs:1-58. In another embodiment, the sequences are sequence variants as further described herein.

Preferably, the cancer modulatory protein is a fragment of about 14-24 amino acids long. More preferably the fragment is a soluble fragment. Preferably, the fragment includes a non-transmembrane region. In a preferred embodiment, the fragment has an N-terminal Cys to aid in solubility. In one embodiment, the C-terminus of the fragment is kept as a free acid and the N-terminus is a free amine to aid in coupling, e.g., to cysteine.

In one embodiment the cancer proteins are conjugated to an immunogenic agent as discussed herein. In one embodiment the cancer protein is conjugated to BSA.

Measurements of cancer polypeptide activity, or of cancer or the cancer phenotype can be performed using a variety of assays. For example, the effects of the test compounds upon the function of the cancer polypeptides can be measured by examining parameters described above. A suitable physiological change that affects activity can be used to assess the influence of a test compound on the polypeptides of this invention. When the functional consequences are determined using intact cells or animals, one can also measure a variety of effects such as, in the case of cancer associated with tumors, tumor growth, tumor metastasis, neovascularization, hormone release, transcriptional changes to both known and uncharacterized genetic markers (e.g., northern blots), changes in cell metabolism such as cell growth or pH changes, and changes in intracellular second messengers such as cGMP.

In the assays of the invention, mammalian cancer polypeptide is typically used, e.g., mouse, preferably human.

Assays to identify compounds with modulating activity can be performed in vitro. For example, a cancer polypeptide is first contacted with a potential modulator and incubated for a suitable amount of time, e.g., from 0.5-48 hours. In one embodiment, the cancer polypeptide levels are determined in vitro by measuring the level of protein or mRNA. The level of protein is typically measured using immunoassays such as western blotting, ELISA, and the like with an antibody that selectively binds to the cancer polypeptide or a fragment thereof. For measurement of mRNA, amplification, e.g., using PCR, LCR, or hybridization assays, e.g., northern hybridization, RNAse protection, dot blotting, are preferred. The level of protein or mRNA is typically detected using directly or indirectly labeled detection agents, e.g., fluorescently or radioactively labeled nucleic acids, radioactively or enzymatically labeled antibodies, and the like, as described herein.

Alternatively, a reporter gene system can be devised using a cancer protein promoter operably linked to a reporter gene such as luciferase, green fluorescent protein, CAT, or β-gal. The reporter construct is typically transfected into a cell. After treatment with a potential modulator, the amount of reporter gene transcription, translation, or activity is measured according to standard techniques.

In a preferred embodiment, as outlined above, screens may be done on individual genes and gene products (proteins). That is, having identified a particular differentially expressed gene as important in a particular state, screening of modulators of the expression of the gene or the gene product itself can be done. The gene products of differentially expressed genes are sometimes referred to herein as "cancer proteins." The cancer protein may be a fragment, or alternatively, the full length protein to a fragment shown herein.

In one embodiment, screening for modulators of expression of specific genes is performed. Typically, the expression of only one or a few genes are evaluated. In another embodiment, screens are designed to first find compounds that bind to differentially expressed proteins. These compounds are then evaluated for the ability to modulate differentially expressed activity. Moreover, once initial candidate compounds are identified, variants can be further screened to better evaluate structure activity relationships.

In a preferred embodiment, binding assays are done. In general, purified or isolated gene product is used; that is, the gene products of one or more differentially expressed nucleic acids are made. For example, antibodies are generated to the protein gene products,

and standard immunoassays are run to determine the amount of protein present. Alternatively, cells comprising the cancer proteins can be used in the assays.

Thus, in a preferred embodiment, the methods comprise combining a cancer protein and a candidate compound, and determining the binding of the compound to the cancer protein. Preferred embodiments utilize the human cancer protein, although other mammalian proteins may also be used, e.g., for the development of animal models of human disease. In some embodiments, as outlined herein, variant or derivative cancer proteins may be used.

Generally, in a preferred embodiment of the methods herein, the cancer protein or the candidate agent is non-diffusably bound to an insoluble support, preferably having isolated sample receiving areas (e.g., a microtiter plate, an array, etc.). The insoluble supports may be made of a composition to which the compositions can be bound, is readily separated from soluble material, and is otherwise compatible with the overall method of screening. The surface of such supports may be solid or porous and of a convenient shape. Examples of suitable insoluble supports include microtiter plates, arrays, membranes, and beads. These are typically made of glass, plastic (e.g., polystyrene), polysaccharides, nylon or nitrocellulose, teflon™, etc. Microtiter plates and arrays are especially convenient because a large number of assays can be carried out simultaneously, using small amounts of reagents and samples. The particular manner of binding of the composition is typically not crucial so long as it is compatible with the reagents and overall methods of the invention, maintains the activity of the composition, and is nondiffusable. Preferred methods of binding include the use of antibodies (which do not sterically block either the ligand binding site or activation sequence when the protein is bound to the support), direct binding to "sticky" or ionic supports, chemical crosslinking, the synthesis of the protein or agent on the surface, etc. Following binding of the protein or agent, excess unbound material is removed by washing. The sample receiving areas may then be blocked through incubation with bovine serum albumin (BSA), casein, or other innocuous protein or other moiety.

In a preferred embodiment, the cancer protein is bound to the support, and a test compound is added to the assay. Alternatively, the candidate agent is bound to the support and the cancer protein is added. Novel binding agents include specific antibodies, non-natural binding agents identified in screens of chemical libraries, peptide analogs, etc. Of particular interest are screening assays for agents that have a low toxicity for human cells. A wide variety of assays may be used for this purpose, including labeled in vitro protein-

protein binding assays, electrophoretic mobility shift assays, immunoassays for protein binding, functional assays (phosphorylation assays, etc.), and the like.

The determination of the binding of the test modulating compound to the cancer protein may be done in a number of ways. In a preferred embodiment, the compound is labeled, and binding determined directly, e.g., by attaching all or a portion of the cancer protein to a solid support, adding a labeled candidate agent (e.g., a fluorescent label), washing off excess reagent, and determining whether the label is present on the solid support. Various blocking and washing steps may be utilized as appropriate.

In some embodiments, only one of the components is labeled, e.g., the proteins (or proteinaceous candidate compounds) can be labeled. Alternatively, more than one component can be labeled with different labels, e.g., ^{125}I for the proteins and a fluorophor for the compound. Proximity reagents, e.g., quenching or energy transfer reagents are also useful.

In one embodiment, the binding of the test compound is determined by competitive binding assay. The competitor may be a binding moiety known to bind to the target molecule (e.g., a cancer protein), such as an antibody, peptide, binding partner, ligand, etc. Under certain circumstances, there may be competitive binding between the compound and the binding moiety, with the binding moiety displacing the compound. In one embodiment, the test compound is labeled. Either the compound, or the competitor, or both, is added first to the protein for a time sufficient to allow binding, if present. Incubations may be performed at a temperature which facilitates optimal activity, typically between about 4-40° C. Incubation periods are typically optimized, e.g., to facilitate rapid high throughput screening. Typically between 0.1-1 hour will be sufficient. Excess reagent is generally removed or washed away. The second component is then added, and the presence or absence of the labeled component is followed, to indicate binding.

In a preferred embodiment, the competitor is added first, followed by a test compound. Displacement of the competitor is an indication that the test compound is binding to the cancer protein and thus is capable of binding to, and potentially modulating, the activity of the cancer protein. In this embodiment, either component can be labeled. Thus, e.g., if the competitor is labeled, the presence of label in the wash solution indicates displacement by the agent. Alternatively, if the test compound is labeled, the presence of the label on the support indicates displacement.

In an alternative embodiment, the test compound is added first, with incubation and washing, followed by the competitor. The absence of binding by the competitor may indicate that the test compound is bound to the cancer protein with a higher affinity. Thus, if the test compound is labeled, the presence of the label on the support, coupled with a lack of competitor binding, may indicate that the test compound is capable of binding to the cancer protein.

In a preferred embodiment, the methods comprise differential screening to identity agents that are capable of modulating the activity of the cancer proteins. In one embodiment, the methods comprise combining a cancer protein and a competitor in a first sample. A second sample comprises a test compound, a cancer protein, and a competitor. The binding of the competitor is determined for both samples, and a change, or difference in binding between the two samples indicates the presence of an agent capable of binding to the cancer protein and potentially modulating its activity. That is, if the binding of the competitor is different in the second sample relative to the first sample, the agent is capable of binding to the cancer protein.

Alternatively, differential screening is used to identify drug candidates that bind to the native cancer protein, but cannot bind to modified cancer proteins. The structure of the cancer protein may be modeled, and used in rational drug design to synthesize agents that interact with that site. Drug candidates that affect the activity of a cancer protein are also identified by screening drugs for the ability to either enhance or reduce the activity of the protein.

Positive controls and negative controls may be used in the assays. Preferably control and test samples are performed in at least triplicate to obtain statistically significant results. Incubation of all samples is for a time sufficient for the binding of the agent to the protein. Following incubation, samples are washed free of non-specifically bound material and the amount of bound, generally labeled agent determined. For example, where a radiolabel is employed, the samples may be counted in a scintillation counter to determine the amount of bound compound.

A variety of other reagents may be included in the screening assays. These include reagents like salts, neutral proteins, e.g., albumin, detergents, etc., which may be used to facilitate optimal protein-protein binding and/or reduce non-specific or background interactions. Also reagents that otherwise improve the efficiency of the assay, such as

protease inhibitors, nuclease inhibitors, anti-microbial agents, etc., may be used. The mixture of components may be added in an order that provides for the requisite binding.

In a preferred embodiment, the invention provides methods for screening for a compound capable of modulating the activity of a cancer protein. The methods comprise adding a test compound, as defined above, to a cell comprising cancer proteins. Preferred cell types include almost any cell. The cells contain a recombinant nucleic acid that encodes a cancer protein. In a preferred embodiment, a library of candidate agents are tested on a plurality of cells.

In one aspect, the assays are evaluated in the presence or absence or previous or subsequent exposure of physiological signals, e.g., hormones, antibodies, peptides, antigens, cytokines, growth factors, action potentials, pharmacological agents including chemotherapeutics, radiation, carcinogenics, or other cells (e.g., cell-cell contacts). In another example, the determinations are determined at different stages of the cell cycle process.

In this way, compounds that modulate cancer agents are identified. Compounds with pharmacological activity are able to enhance or interfere with the activity of the cancer protein. Once identified, similar structures are evaluated to identify critical structural feature of the compound.

In one embodiment, a method of inhibiting cancer cell division is provided. The method comprises administration of a cancer inhibitor. In another embodiment, a method of inhibiting cancer is provided. The method may comprise administration of a cancer inhibitor. In a further embodiment, methods of treating cells or individuals with cancer are provided, e.g., comprising administration of a cancer inhibitor.

In one embodiment, a cancer inhibitor is an antibody as discussed above. In another embodiment, the cancer inhibitor is an antisense molecule.

A variety of cell growth, proliferation, viability, and metastasis assays are available, as described below.

Soft agar growth or colony formation in suspension

Normal cells require a solid substrate to attach and grow. When the cells are transformed, they lose this phenotype and grow detached from the substrate. For example, transformed cells can grow in stirred suspension culture or suspended in semi-solid media, such as semi-solid or soft agar. The transformed cells, when transfected with tumor suppressor genes, regenerate normal phenotype and require a solid substrate to attach and

grow. Soft agar growth or colony formation in suspension assays can be used to identify modulators of cancer sequences, which when expressed in host cells, inhibit abnormal cellular proliferation and transformation. A therapeutic compound would reduce or eliminate the host cells' ability to grow in stirred suspension culture or suspended in semi-solid media, such as semi-solid or soft.

Techniques for soft agar growth or colony formation in suspension assays are described, e.g., in Freshney (1998) Culture of Animal Cells: A Manual of Basic Technique (3d ed.) Wiley-Liss; Freshney (2000) Culture of Animal Cells: A Manual of Basic Technique (4th ed.) Wiley-Liss; and Garkavtsev, et al. (1996) Nature Genet. 14:415-20.

Contact inhibition and density limitation of growth

Normal cells typically grow in a flat and organized pattern in a petri dish until they touch other cells. When the cells touch one another, they are contact inhibited and stop growing. When cells are transformed, however, the cells are not contact inhibited and continue to grow to high densities in disorganized foci. Thus, the transformed cells grow to a higher saturation density than normal cells. This can be detected morphologically by the formation of a disoriented monolayer of cells or rounded cells in foci within the regular pattern of normal surrounding cells. Alternatively, labeling index with (^{3}H)-thymidine at saturation density can be used to measure density limitation of growth. See Freshney (2000), supra. The transformed cells, when transfected with tumor suppressor genes, regenerate a normal phenotype and become contact inhibited and would grow to a lower density.

In this assay, labeling index with (^{3}H)-thymidine at saturation density is a preferred method of measuring density limitation of growth. Transformed host cells are transfected with a cancer-associated sequence and are grown for 24 hours at saturation density in non-limiting medium conditions. The percentage of cells labeling with (^{3}H)-thymidine is determined autoradiographically. See, Freshney (1998), supra.

Growth factor or serum dependence

Transformed cells typically have a lower serum dependence than their normal counterparts (see, e.g., Temin (1966) J. Natl. Cancer Insti. 37:167-175; Eagle, et al.(1970) J. Exp. Med. 131:836-879); Freshney, supra. This is in part due to release of various growth factors by the transformed cells. Growth factor or serum dependence of transformed host cells can be compared with that of control.

Tumor specific markers levels

Tumor cells release an increased amount of certain factors (hereinafter "tumor specific markers") than their normal counterparts. For example, plasminogen activator (PA) is released from human glioma at a higher level than from normal brain cells (see, e.g., Gullino "Angiogenesis, tumor vascularization, and potential interference with tumor growth" pp. 178-184 in Mihich (ed. 1985) Biological Responses in Cancer Plenum. Similarly, tumor angiogenesis factor (TAF) is released at a higher level in tumor cells than their normal counterparts. See, e.g., Folkman (1992) Sem. Cancer Biol. 3:89-96.

Various techniques which measure the release of these factors are described in Freshney (1998), supra. Also, see, Unkeless, et al. (1974) J. Biol. Chem. 249:4295-4305; Strickland and Beers (1976) J. Biol. Chem. 251:5694-5702; Whur, et al. (1980) Br. J. Cancer 42:305-312; Gullino "Angiogenesis, tumor vascularization, and potential interference with tumor growth" pp. 178-184 in Mihich (ed. 1985) Biological Responses in Cancer Plenum; Freshney (1985) Anticancer Res. 5:111-130.

Invasiveness into Matrigel

The degree of invasiveness into Matrigel or some other extracellular matrix constituent can be used as an assay to identify compounds that modulate cancer-associated sequences. Tumor cells exhibit a good correlation between malignancy and invasiveness of cells into Matrigel or some other extracellular matrix constituent. In this assay, tumorigenic cells are typically used as host cells. Expression of a tumor suppressor gene in these host cells would decrease invasiveness of the host cells.

Techniques described in Freshney (1994), supra, can be used. Briefly, the level of invasion of host cells can be measured by using filters coated with Matrigel or some other extracellular matrix constituent. Penetration into the gel, or through to the distal side of the filter, is rated as invasiveness, and rated histologically by number of cells and distance moved, or by prelabeling the cells with ^{125}I and counting the radioactivity on the distal side of the filter or bottom of the dish. See, e.g., Freshney (1984), supra.

Tumor growth in vivo

Effects of cancer-associated sequences on cell growth can be tested in transgenic or immune-suppressed mice. Knock-out transgenic mice can be made, in which the cancer gene is disrupted or in which a cancer gene is inserted. Knock-out transgenic mice can be made by insertion of a marker gene or other heterologous gene into the endogenous cancer gene site in the mouse genome via homologous recombination. Such mice can also be

made by substituting the endogenous cancer gene with a mutated version of the cancer gene, or by mutating the endogenous cancer gene, e.g., by exposure to carcinogens.

A DNA construct is introduced into the nuclei of embryonic stem cells. Cells containing the newly engineered genetic lesion are injected into a host mouse embryo, which is re-implanted into a recipient female. Some of these embryos develop into chimeric mice that possess germ cells partially derived from the mutant cell line. Therefore, by breeding the chimeric mice it is possible to obtain a new line of mice containing the introduced genetic lesion (see, e.g., Capecchi, et al. (1989) Science 244:1288-1292). Chimeric targeted mice can be derived according to Hogan, et al. (1988) Manipulating the Mouse Embryo: A Laboratory Manual CSH Press; and Robertson (ed. 1987) Teratocarcinomas and Embryonic Stem Cells: A Practical Approach IRL Press, Washington, D.C.

Alternatively, various immune-suppressed or immune-deficient host animals can be used. For example, genetically athymic "nude" mouse (see, e.g., Giovanella, et al. (1974) J. Natl. Cancer Inst. 52:921-930), a SCID mouse, a thymectomized mouse, or an irradiated mouse (see, e.g., Bradley, et al. (1978) Br. J. Cancer 38:263-272; Selby, et al. (1980) Br. J. Cancer 41:52-61) can be used as a host. Transplantable tumor cells (typically about 10^6 cells) injected into isogenic hosts will produce invasive tumors in a high proportions of cases, while normal cells of similar origin will not. In hosts which developed invasive tumors, cells expressing a cancer-associated sequences are injected subcutaneously. After a suitable length of time, preferably about 4-8 weeks, tumor growth is measured (e.g., by volume or by its two largest dimensions) and compared to the control. Tumors that have statistically significant reduction (using, e.g., Student's T test) are said to have inhibited growth.

Polynucleotide modulators of cancer

Antisense and RNAi Polynucleotides

In certain embodiments, the activity of a cancer-associated protein is down-regulated, or entirely inhibited, by the use of an inhibitory or antisense polynucleotide, e.g., a nucleic acid complementary to, and which can preferably hybridize specifically to, a coding mRNA nucleic acid sequence, e.g., a cancer protein mRNA, or a subsequence thereof. Binding of the antisense polynucleotide to the mRNA reduces the translation and/or stability of the mRNA.

In the context of this invention, antisense polynucleotides can comprise naturally-occurring nucleotides, or synthetic species formed from naturally-occurring subunits or their close homologs. Antisense polynucleotides may also have altered sugar moieties or inter-sugar linkages. Exemplary among these are the phosphorothioate and other sulfur containing species. Analogs are comprehended by this invention so long as they function effectively to hybridize with the cancer protein mRNA. See, e.g., Isis Pharmaceuticals, Carlsbad, CA; Sequitor, Inc., Natick, MA.

Such antisense polynucleotides can readily be synthesized using recombinant means, or can be synthesized in vitro. Equipment for such synthesis is sold by several vendors, including Applied Biosystems. The preparation of other oligonucleotides such as phosphorothioates and alkylated derivatives is also well known.

Antisense molecules as used herein include antisense or sense oligonucleotides. Sense oligonucleotides can, e.g., be employed to block transcription by binding to the anti-sense strand. The antisense and sense oligonucleotide comprise a single-stranded nucleic acid sequence (either RNA or DNA) capable of binding to target mRNA (sense) or DNA (antisense) sequences for cancer molecules. A preferred antisense molecule is for a cancer sequence in the Table 2 or the attached listing of SEQ ID NOs:1-116, or for a ligand or activator thereof. Antisense or sense oligonucleotides, according to the present invention, comprise a fragment generally at least about 14 nucleotides, preferably from about 14-30 nucleotides. The ability to derive an antisense or a sense oligonucleotide, based upon a cDNA sequence encoding a given protein is described in, e.g., Stein and Cohen (1988) Cancer Res. 48:2659-2668; and van der Krol, et al. (1988) BioTechniques 6:958-976.

RNA interference is a mechanism to suppress gene expression in a sequence specific manner. See, e.g., Brumelkamp, et al. (2002) Sciencexpress (21March2002); Sharp (1999) Genes Dev. 13:139-141; and Cathew (2001) Curr. Op. Cell Biol. 13:244-248. In mammalian cells, short, e.g., 21 nt, double stranded small interfering RNAs (siRNA) have been shown to be effective at inducing an RNAi response. See, e.g., Elbashir, et al. (2001) Nature 411:494-498. The mechanism may be used to downregulate expression levels of identified genes, e.g., treatment of or validation of relevance to disease.

Ribozymes

In addition to antisense polynucleotides, ribozymes can be used to target and inhibit transcription of cancer-associated nucleotide sequences. A ribozyme is an RNA molecule that catalytically cleaves other RNA molecules. Different kinds of ribozymes have been

described, including group I ribozymes, hammerhead ribozymes, hairpin ribozymes, RNase P, and axhead ribozymes (see, e.g., Castanotto, et al. (1994) Adv. in Pharmacology 25: 289-317 for a general review of the properties of different ribozymes).

The general features of hairpin ribozymes are described, e.g., in Hampel, et al. (1990) Nucl. Acids Res. 18:299-304; European Patent Publication No. 0 360 257; U.S. Patent No. 5,254,678. Methods of preparation are described in, e.g., WO 94/26877; Ojwang, et al. (1993) Proc. Natl. Acad. Sci. USA 90:6340-6344; Yamada, et al. (1994) Human Gene Therapy 1:39-45; Leavitt, et al.(1995) Proc. Natl. Acad. Sci. USA 92:699-703; Leavitt, et al. (1994) Human Gene Therapy 5:1151-120; and Yamada, et al. (1994) Virology 205: 121-126.

Polynucleotide modulators of cancer may be introduced into a cell containing the target nucleotide sequence by formation of a conjugate with a ligand binding molecule, as described in WO 91/04753. Suitable ligand binding molecules include, but are not limited to, cell surface receptors, growth factors, other cytokines, or other ligands that bind to cell surface receptors. Preferably, conjugation of the ligand binding molecule does not substantially interfere with the ability of the ligand binding molecule to bind to its corresponding molecule or receptor, or block entry of the sense or antisense oligonucleotide or its conjugated version into the cell. Alternatively, a polynucleotide modulator of cancer may be introduced into a cell containing the target nucleic acid sequence, e.g., by formation of an polynucleotide-lipid complex, as described in WO 90/10448. It is understood that the use of antisense molecules or knock out and knock in models may also be used in screening assays as discussed above, in addition to methods of treatment.

Thus, in one embodiment, methods of modulating cancer in cells or organisms are provided. In one embodiment, the methods comprise administering to a cell an anti-cancer antibody that reduces or eliminates the biological activity of an endogenous cancer protein. Alternatively, the methods comprise administering to a cell or organism a recombinant nucleic acid encoding a cancer protein. This may be accomplished in any number of ways. In a preferred embodiment, e.g., when the cancer sequence is down-regulated in cancer, such state may be reversed by increasing the amount of cancer gene product in the cell. This can be accomplished, e.g., by overexpressing the endogenous cancer gene or administering a gene encoding the cancer sequence, using known gene-therapy techniques. In a preferred embodiment, the gene therapy techniques include the incorporation of the exogenous gene using enhanced homologous recombination (EHR), e.g., as described in

PCT/US93/0386. Alternatively, e.g., when the cancer sequence is up-regulated in cancer, the activity of the endogenous cancer gene is decreased, e.g., by the administration of a cancer antisense or other inhibitor, e.g., RNAi.

In one embodiment, the cancer proteins of the present invention may be used to generate polyclonal and monoclonal antibodies to cancer proteins. Similarly, the cancer proteins can be coupled, using standard technology, to affinity chromatography columns. These columns may then be used to purify cancer antibodies useful for production, diagnostic, or therapeutic purposes. In a preferred embodiment, the antibodies are generated to epitopes unique to a cancer protein; that is, the antibodies show little or no cross-reactivity to other proteins. The cancer antibodies may be coupled to standard affinity chromatography columns and used to purify cancer proteins. The antibodies may also be used as blocking polypeptides, as outlined above, since they will specifically bind to the cancer protein.

Methods of identifying variant cancer-associated sequences

Without being bound by theory, expression of various cancer sequences is correlated with cancer. Accordingly, disorders based on mutant or variant cancer genes may be determined. In one embodiment, the invention provides methods for identifying cells containing variant cancer genes, e.g., determining all or part of the sequence of at least one endogenous cancer gene in a cell. In a preferred embodiment, the invention provides methods of identifying the cancer genotype of an individual, e.g., determining all or part of the sequence of at least one cancer gene of the individual. This is generally done in at least one tissue of the individual, and may include the evaluation of a number of tissues or different samples of the same tissue. The method may include comparing the sequence of the sequenced cancer gene to a known cancer gene, e.g., a wild-type gene.

The sequence of all or part of the cancer gene can then be compared to the sequence of a known cancer gene to determine if any differences exist. This can be done using known homology programs, such as Bestfit, etc. In a preferred embodiment, the presence of a difference in the sequence between the cancer gene of the patient and the known cancer gene correlates with a disease state or a propensity for a disease state, as outlined herein.

In a preferred embodiment, the cancer genes are used as probes to determine the number of copies of the cancer gene in the genome.

In another preferred embodiment, the cancer genes are used as probes to determine the chromosomal localization of the cancer genes. Information such as chromosomal

localization finds use in providing a diagnosis or prognosis in particular when chromosomal abnormalities such as translocations, and the like are identified in the cancer gene locus.

Administration of pharmaceutical and vaccine compositions

In one embodiment, a therapeutically effective dose of a cancer protein or modulator thereof, is administered to a patient. By "therapeutically effective dose" herein is meant a dose that produces effects for which it is administered. The exact dose will depend on the purpose of the treatment, and will be ascertainable using known techniques. See, e.g., Ansel, et al. (1999) Pharmaceutical Dosage Forms and Drug Delivery Lippincott; Lieberman (1992) Pharmaceutical Dosage Forms (vols. 1-3) Dekker, ISBN 0824770846, 082476918X, 0824712692, 0824716981; Lloyd (1999) The Art, Science and Technology of Pharmaceutical Compounding Amer. Pharmaceut. Assn.; and Pickar (1998) Dosage Calculations Thomson. Adjustments for cancer degradation, systemic versus localized delivery, and rate of new protease synthesis, as well as the age, body weight, general health, sex, diet, time of administration, drug interaction, and the severity of the condition may be necessary. U.S. Patent Application No. 09/687,576, further discloses the use of compositions and methods of diagnosis and treatment in cancer.

A "patient" for the purposes of the present invention includes both humans and other animals, particularly mammals. Thus the methods are applicable to both human therapy and veterinary applications. In the preferred embodiment the patient is a mammal, preferably a primate, and in the most preferred embodiment the patient is human.

The administration of the cancer proteins and modulators thereof of the present invention can be done in a variety of ways, including, but not limited to, orally, subcutaneously, intravenously, intranasally, transdermally, intraperitoneally, intramuscularly, intrapulmonary, vaginally, rectally, or intraocularly. In some instances, e.g., in the treatment of wounds and inflammation, the cancer proteins and modulators may be directly applied as a solution or spray.

The pharmaceutical compositions of the present invention comprise a cancer protein in a form suitable for administration to a patient. In the preferred embodiment, the pharmaceutical compositions are in a water soluble form, such as being present as pharmaceutically acceptable salts, which is meant to include both acid and base addition salts. "Pharmaceutically acceptable acid addition salt" refers to those salts that retain the biological effectiveness of the free bases and that are not biologically or otherwise undesirable, formed with inorganic acids such as hydrochloric acid, hydrobromic acid,

sulfuric acid, nitric acid, phosphoric acid, and the like, and organic acids such as acetic acid, propionic acid, glycolic acid, pyruvic acid, oxalic acid, maleic acid, malonic acid, succinic acid, fumaric acid, tartaric acid, citric acid, benzoic acid, cinnamic acid, mandelic acid, methanesulfonic acid, ethanesulfonic acid, p-toluenesulfonic acid, salicylic acid, and the like. "Pharmaceutically acceptable base addition salts" include those derived from inorganic bases such as sodium, potassium, lithium, ammonium, calcium, magnesium, iron, zinc, copper, manganese, aluminum salts, and the like. Particularly preferred are the ammonium, potassium, sodium, calcium, and magnesium salts. Salts derived from pharmaceutically acceptable organic non-toxic bases include salts of primary, secondary, and tertiary amines, substituted amines including naturally occurring substituted amines, cyclic amines and basic ion exchange resins, such as isopropylamine, trimethylamine, diethylamine, triethylamine, tripropylamine, and ethanolamine.

The pharmaceutical compositions may also include one or more of the following: carrier proteins such as serum albumin; buffers; fillers such as microcrystalline cellulose, lactose, corn and other starches; binding agents; sweeteners and other flavoring agents; coloring agents; and polyethylene glycol.

The pharmaceutical compositions can be administered in a variety of unit dosage forms depending upon the method of administration. For example, unit dosage forms suitable for oral administration include, but are not limited to, powder, tablets, pills, capsules and lozenges. It is recognized that cancer protein modulators (e.g., antibodies, antisense constructs, ribozymes, small organic molecules, etc.) when administered orally, should be protected from digestion. This is typically accomplished either by complexing the molecule(s) with a composition to render it resistant to acidic and enzymatic hydrolysis, or by packaging the molecule(s) in an appropriately resistant carrier, such as a liposome or a protection barrier. Means of protecting agents from digestion are available.

The compositions for administration will commonly comprise a cancer protein modulator dissolved in a pharmaceutically acceptable carrier, preferably an aqueous carrier. A variety of aqueous carriers can be used, e.g., buffered saline and the like. These solutions are sterile and generally free of undesirable matter. These compositions may be sterilized by conventional, well known sterilization techniques. The compositions may contain pharmaceutically acceptable auxiliary substances as required to approximate physiological conditions such as pH adjusting and buffering agents, toxicity adjusting agents, and the like, e.g., sodium acetate, sodium chloride, potassium chloride, calcium chloride, sodium lactate,

and the like. The concentration of active agent in these formulations can vary widely, and will be selected primarily based on fluid volumes, viscosities, body weight, and the like in accordance with the particular mode of administration selected and the patient's needs (e.g., (1980) Remington's Pharmaceutical Science (18th ed.) Mack, and Hardman and Limbird (eds. 2001) Goodman and Gilman: The Pharmacological Basis of Therapeutics (10th ed.) McGraw-Hill.

Thus, a typical pharmaceutical composition for intravenous administration would be about 0.1 to 10 mg per patient per day. Dosages from 0.1 up to about 100 mg per patient per day may be used, particularly when the drug is administered to a secluded site and not into the blood stream, such as into a body cavity or into a lumen of an organ. Substantially higher dosages are possible in topical administration. Actual methods for preparing parenterally administrable compositions will be known or apparent.

The compositions containing modulators of cancer proteins can be administered for therapeutic or prophylactic treatments. In therapeutic applications, compositions are administered to a patient suffering from a disease (e.g., a cancer) in an amount sufficient to cure or at least partially arrest the disease and its complications. An amount adequate to accomplish this is defined as a "therapeutically effective dose." Amounts effective for this use will depend upon the severity of the disease and the general state of the patient's health. Single or multiple administrations of the compositions may be administered depending on the dosage and frequency as required and tolerated by the patient. In any event, the composition should provide a sufficient quantity of the agents of this invention to effectively treat the patient. An amount of modulator that is capable of preventing or slowing the development of cancer in a mammal is referred to as a "prophylactically effective dose." The particular dose required for a prophylactic treatment will depend upon the medical condition and history of the mammal, the particular cancer being prevented, as well as other factors such as age, weight, gender, administration route, efficiency, etc. Such prophylactic treatments may be used, e.g., in a mammal who has previously had cancer to prevent a recurrence of the cancer, or in a mammal who is suspected of having a significant likelihood of developing cancer based, at least in part, upon gene expression profiles. Vaccine strategies may be used, in either a DNA vaccine form, or protein vaccine.

It will be appreciated that the present cancer protein-modulating compounds can be administered alone or in combination with additional cancer modulating compounds or with other therapeutic agent, e.g., other anti-cancer agents or treatments.

In numerous embodiments, one or more nucleic acids, e.g., polynucleotides comprising nucleic acid sequences set forth in Table 2 or the attached listing of SEQ ID NOs:1-58, such as RNAi, antisense polynucleotides or ribozymes, will be introduced into cells, in vitro or in vivo. The present invention provides methods, reagents, vectors, and cells useful for expression of cancer-associated polypeptides and nucleic acids using in vitro (cell-free), ex vivo or in vivo (cell or organism-based) recombinant expression systems.

The particular procedure used to introduce the nucleic acids into a host cell for expression of a protein or nucleic acid is application specific. Many procedures for introducing foreign nucleotide sequences into host cells may be used. These include the use of calcium phosphate transfection, spheroplasts, electroporation, liposomes, microinjection, plasma vectors, viral vectors, and other well known methods for introducing cloned genomic DNA, cDNA, synthetic DNA, or other foreign genetic material into a host cell (see, e.g., Berger and Kimmel (1987) Guide to Molecular Cloning Techniques from Methods in Enzymology (vol. 152) Academic Press; Ausubel, et al. (eds. 1999 and supplements) Current Protocols Lippincott; and Sambrook, et al. (2001) Molecular Cloning: A Laboratory Manual (3d ed., Vol. 1-3) CSH Press.

In a preferred embodiment, cancer proteins and modulators are administered as therapeutic agents, and can be formulated as outlined above. Similarly, cancer genes (including both the full-length sequence, partial sequences, or regulatory sequences of the cancer coding regions) can be administered in a gene therapy application. These cancer genes can include inhibitory applications, e.g., as inhibitory RNA, gene therapy (e.g., for incorporation into the genome), or antisense compositions.

Cancer polypeptides and polynucleotides can also be administered as vaccine compositions to stimulate HTL, CTL, and antibody responses. Such vaccine compositions can include, e.g., lipidated peptides (see, e.g.,Vitiello, et al. (1995) J. Clin. Invest. 95:341-349), peptide compositions encapsulated in poly(DL-lactide-co-glycolide) ("PLG") microspheres (see, e.g., Eldridge, et al. (1991) Molec. Immunol. 28:287-294,; Alonso, et al. (1994) Vaccine 12:299-306; Jones, et al. (1995) Vaccine 13:675-681), peptide compositions contained in immune stimulating complexes (ISCOMS) (see, e.g., Takahashi, et al. (1990) Nature 344:873-875; Hu, et al. (1998) Clin Exp Immunol. 113:235-243), multiple antigen peptide systems (MAPs) (see, e.g., Tam (1988) Proc. Natl. Acad. Sci. USA 85:5409-5413; Tam (1996) J. Immunol. Methods 196:17-32), peptides formulated as multivalent peptides; peptides for use in ballistic delivery systems, typically crystallized peptides, viral delivery

vectors (Perkus, et al., p. 379, in Kaufmann (ed. 1996) Concepts in Vaccine Development de Gruyter; Chakrabarti, et al. (1986) Nature 320:535-537; Hu, et al. (1986) Nature 320:537-540; Kieny, et al. (1986) Bio/Technology 4:790-795; Top, et al. (1971) J. Infect. Dis. 124:148-154; Chanda, et al. (1990) Virology 175:535-547), particles of viral or synthetic origin (see, e.g., Kofler, et al. (1996) J. Immunol. Methods 192:25-35; Eldridge, et al. (1993) Sem. Hematol. 30:16-24; Falo, et al. (1995) Nature Med. 1:649-653), adjuvants (Warren, et al. (1986) Annu. Rev. Immunol. 4:369-388; Gupta, et al. (1993) Vaccine 11:293-306), liposomes (Reddy, et al. (1992) J. Immunol. 148:1585-1589; Rock (1996) Immunol. Today 17:131-137), or, naked or particle absorbed cDNA (Ulmer, et al. (1993) Science 259:1745-1749; Robinson, et al. (1993) Vaccine 11:957-960; Shiver, et al., p 423, in Kaufmann (ed. 1996) Concepts in Vaccine Development de Gruyter; Cease and Berzofsky (1994) Annu. Rev. Immunol. 12:923-989; and Eldridge, et al. (1993) Sem. Hematol. 30:16-24). Toxin-targeted delivery technologies, also known as receptor mediated targeting, such as those of Avant Immunotherapeutics, Inc. (Needham, Massachusetts) may also be used.

Vaccine compositions often include adjuvants. Many adjuvants contain a substance designed to protect the antigen from rapid catabolism, such as aluminum hydroxide or mineral oil, and a stimulator of immune responses, such as lipid A, Bortadella pertussis, or Mycobacterium tuberculosis derived proteins. Certain adjuvants are commercially available as, e.g., Freund's Incomplete Adjuvant and Complete Adjuvant (Difco Laboratories, Detroit, MI); Merck Adjuvant 65 (Merck and Company, Inc., Rahway, NJ); AS-2 (SmithKline Beecham, Philadelphia, PA); aluminum salts such as aluminum hydroxide gel (alum) or aluminum phosphate; salts of calcium, iron, or zinc; an insoluble suspension of acylated tyrosine; acylated sugars; cationically or anionically derivatized polysaccharides; polyphosphazenes; biodegradable microspheres; monophosphoryl lipid A and quil A. Cytokines, such as GM-CSF, interleukin-2, -7, -12, and other like growth factors, may also be used as adjuvants.

Vaccines can be administered as nucleic acid compositions wherein DNA or RNA encoding one or more of the polypeptides, or a fragment thereof, is administered to a patient. This approach is described, for instance, in Wolff et. al. (1990) Science 247:1465-1468, as well as U.S. Patent Nos. 5,580,859; 5,589,466; 5,804,566; 5,739,118; 5,736,524; 5,679,647; WO 98/04720; and in more detail below. Examples of DNA-based delivery technologies include "naked DNA", facilitated (bupivicaine, polymers, peptide-mediated)

delivery, cationic lipid complexes, and particle-mediated ("gene gun") or pressure-mediated delivery (see, e.g., U.S. Patent No. 5,922,687).

For therapeutic or prophylactic immunization purposes, the peptides of the invention can be expressed by viral or bacterial vectors. Examples of expression vectors include attenuated viral hosts, such as vaccinia or fowlpox. This approach involves the use of vaccinia virus, e.g., as a vector to express nucleotide sequences that encode cancer polypeptides or polypeptide fragments. Upon introduction into a host, the recombinant vaccinia virus expresses the immunogenic peptide, and thereby elicits an immune response. Vaccinia vectors and methods useful in immunization protocols are described in, e.g., U.S. Patent No. 4,722,848. Another vector is BCG (Bacille Calmette Guerin). BCG vectors are described in Stover, et al. (1991) Nature 351:456-460. A wide variety of other vectors are availablel for therapeutic administration or immunization, e.g., adeno and adeno-associated virus vectors, retroviral vectors, Salmonella typhi vectors, detoxified anthrax toxin vectors, and the like. See, e.g., Shata, et al. (2000) Mol Med Today 6:66-71; Shedlock, et al. (2000) J. Leukoc. Biol. 68:793-806; Hipp, et al. (2000) In Vivo 14:571-85.

Methods for the use of genes as DNA vaccines are well known, and include placing a cancer gene or portion of a cancer gene under the control of a regulatable promoter or a tissue-specific promoter for expression in a cancer patient. The cancer gene used for DNA vaccines can encode full-length cancer proteins, but more preferably encodes portions of the cancer proteins including peptides derived from the cancer protein. In one embodiment, a patient is immunized with a DNA vaccine comprising a plurality of nucleotide sequences derived from a cancer gene. For example, cancer-associated genes or sequence encoding subfragments of a cancer protein are introduced into expression vectors and tested for their immunogenicity in the context of Class I MHC and an ability to generate cytotoxic T cell responses. This procedure provides for production of cytotoxic T cell responses against cells which present antigen, including intracellular epitopes.

In a preferred embodiment, DNA vaccines include a gene encoding an adjuvant molecule with the DNA vaccine. Such adjuvant molecules include cytokines that increase the immunogenic response to the cancer polypeptide encoded by the DNA vaccine. Additional or alternative adjuvants are available.

In another preferred embodiment, cancer genes find use in generating animal models of cancer. When the cancer gene identified is repressed or diminished in cancer tissue, gene therapy technology, e.g., wherein inhibitory or antisense RNA directed to the cancer gene

will also diminish or repress expression of the gene. Animal models of cancer find use in screening for modulators of a cancer-associated sequence or modulators of cancer. Similarly, transgenic animal technology, including gene knockout technology, e.g., as a result of homologous recombination with an appropriate gene targeting vector, will result in the absence or increased expression of the cancer protein. When desired, tissue-specific expression or knockout of the cancer protein may be necessary.

It is also possible that the cancer protein is overexpressed in cancer. As such, transgenic animals can be generated that overexpress the cancer protein. Depending on the desired expression level, promoters of various strengths can be employed to express the transgene. Also, the number of copies of the integrated transgene can be determined and compared for a determination of the expression level of the transgene. Animals generated by such methods will find use as animal models of cancer and are additionally useful in screening for modulators to treat cancer.

Kits for Use in Diagnostic and/or Prognostic Applications

For use in diagnostic, research, and therapeutic applications suggested above, kits are also provided by the invention. In diagnostic and research applications, such kits may include at least one of the following: assay reagents, buffers, cancer-specific nucleic acids or antibodies, hybridization probes and/or primers, antisense polynucleotides, ribozymes, dominant negative cancer polypeptides or polynucleotides, small molecule inhibitors of cancer-associated sequences etc. A therapeutic product may include sterile saline or another pharmaceutically acceptable emulsion and suspension base.

In addition, the kits may include instructional materials containing instructions (e.g., protocols) for the practice of the methods of this invention. While the instructional materials typically comprise written or printed materials, they are not limited to such. A medium capable of storing such instructions and communicating them to an end user is contemplated by this invention. Such media include, but are not limited to, electronic storage media (e.g., magnetic discs, tapes, cartridges, chips), optical media (e.g., CD ROM), and the like. Such media may include addresses to internet sites that provide such instructional materials.

The present invention also provides for kits for screening for modulators of cancer-associated sequences. Such kits can be prepared from readily available materials and reagents. For example, such kits can comprise one or more of the following materials: a cancer-associated polypeptide or polynucleotide, reaction tubes, and instructions for testing

cancer-associated activity. Optionally, the kit contains biologically active cancer protein. A wide variety of kits and components can be prepared according to the present invention, depending upon the intended user of the kit and the particular needs of the user. Diagnosis would typically involve evaluation of a plurality of genes or products. The genes will typically be selected based on correlations with important parameters in disease which may be identified in historical or outcome data.

EXAMPLES

Example 1: Gene Chip Analysis

Molecular profiles of various normal and cancerous tissues were determined and analyzed using gene chips. RNA was isolated and gene chip analysis was performed as described (Glynne, et al. (2000) Nature 403:672-676; Zhao, et al. (2000) Genes Dev. 14:981-993).

Table 1

Table 1 lists medical conditions, pathologies, abnormalities, or organs affected by disease, referred to in Table 2, for which markers have been identified, and other related medical conditions (including various stages and/or metastases) in which those markers will also be useful, e.g., in therapeutic, diagnostic, prognostic, subsetting, vaccine, and other uses.

Table 1

blood vessels/angiogenesis:	hemangiomas, lymphangiomas, angiosarcoma, lymphangiosarcoma, Kaposi's sarcoma, wound healing, tissue remodeling, psoriasis, ischemic, heart disease, inflammatory diseases (e.g., arthritis, asthma, chronic bronchitis), atherosclerosis, endometriosis, presumed ocular histoplasmosis syndrome, hypoxia, solid tumors, lymphomas, lymphadenitis, lymphangitis, autoimmune diseases (e.g., RA, SLE, juvenile chronic arthritis, pigmented villonodular synovitis, etc.), retinal neovascularization syndromes (e.g., diabetic retinopathy, macular degeneration, presumed ocular histoplasmosis syndrome, etc.), scleritis/conjunctivitis, hypertrophic scars (keloid), birth control, uterine fibroids
bladder:	carcinoma in situ, papillary carcinomas, transitional cell carcinoma, squamous cell carcinoma
bone:	Ewing sarcoma, sarcomas arising from skeletal and extraskeletal connective tissues, including the peripheral nervous system (e.g. chondrosarcoma, osteosarcoma)
brain:	glioblastoma, oligodendroglioma, anablastic astrocytoma, meningioma, medulablastoma, neuroblastoma, ependymoma, schwannoma, craniopharyngioma, pineoblastoma, pineocytoma, neurofibroma, neurofibrosarcoma, malignant peripheral nerve sheath tumors, granular cell tumors, plexosarcoma, ganglioneuroblastoma, neuroepithelioma, neuroma, ganglioneuroma
breast:	ductal carcinoma in situ, lobular carcinoma in situ
cervix:	cancer of the cervix, vagina, or vulva
colon/rectum:	precancerous colorectal disease (e.g., neoplastic polyps (adenomas), familial adenomatous polyposis, ulcerative colitis), colon cancer, e.g., epithelial tumor (e.g., adenocarcinoma, mucinous adenocarcinoma, signet-ring cell adenocarcinoma, squamous cell carcinoma, adenosquamous carcinoma, undifferentiated carcinoma, unclassified carcinoma), carinoid tumor (e.g., argentaffin, nonargentaffin, composite), non-epithelial tumor (e.g., leimyo sarcoma, others), inflammatory bowel disease (e.g., ulcerative colitis, Crohn's disease (granulomatous colitis), dysplasia), rectal cancer, cancer of the anal region (e.g., squamous cell carcinoma, transitional carcinoma, adenocarcinoma, carcinoma, papillary villous carcinoma, mucinous adenocarcinoma, melanoma)
esophagus:	premalignant or predisposing conditions (e.g., esophagitis), squamous cell cancers (e.g., cancers of the head and neck, lung, or cervix), gastrodigestive carcinomas (e.g., cancers of the stomach, colon, or rectum)
fibrosis:	lung fibrosis (idiopathic pulmonary fibrosis, hypersensitivity pneumonitis,interstitial pneumonitis, nonspecific idiopathic pneumonitis), chronic obstructive pulmonary disease (e.g., emphysema, chronic bronchitis), asthma, bronchiectasis, cirrhosis (liver fibrosis), renal fibrosis, scleroderma, wound healing
head and neck:	tumors of the nasal cavity, paranasal sinuses, nasopharynx, oral cavity, oral pharynx, lip, larynx, hypopharynx,

	salivary glands, paragangliomas, esophagus
kidney:	clear cell (nonpapillary) carcinoma, papillary carcinoma, chromophobe renal carcinoma, hypernephroma, adenocarcinoma, sporadic renal carcinomas, hereditary renal carcinomas (von Hippel-Lindau disease), carcinoma of the renal pelvis, ureteral carcinoma, fibroma, papillary adenoma, angiomyolipoma, oncocytoma
leukocytes:	acute lymphoblastic leukemia/lymphoma, chronic lymphocytic leukemia, follicular lymphoma, large B-cell lymphoma, Burkitt lymphoma, plasma cell neoplasms, mantle cell lymphoma, lymphoplasmacytic lymphoma, peripheral T-cell lymphoma, adult T-cell leukemia/lymphoma, Hodgkin disease, acute myelogenous leukemia, chronic myelogenous leukemia, thymic hyperplasia, hairy cell leukemia, malignant transformation, inappropriate activation or abnormalities of leukocytes (e.g., immature, precursor B (pre-B) or precursor T (pre-T) lymphocytes, monocytes, neutrophils, eosinophils, basophils, dendritic cells, lymphoblasts), arthritis, inflammation, leukocytosis, lymphadenitis, lymphangitis, bacteremia, chronic nonspecific lymphadenitis, psoriasis, wound healing
liver:	hepatitis (e.g., types A, B, C), benign epithelial tumors and tumor bile conditions, primary malignant epithelial tumors, primary malignant mesenchymal tumors, tumors of the gallbladder or bile duct
lung:	lung cancer, small cell lung carcinoma (oat cell carcinoma), non-small cell carcinomas (e.g., squamous cell carcinoma, adenocarcinoma, large cell lung carcinoma, carcinoid, granulomatous), fibrosis (idiopathic pulmonary fibrosis, hypersensitivity pneumonitis,interstitial pneumonitis, nonspecific idiopathic pneumonitis), chronic obstructive pulmonary disease (e.g., emphysema, chronic bronchitis), asthma, bronchiectasis, esophageal cancer
ovary:	ovarian carcinoma (e.g., epithelial (serous tumors, mucinous tumors, endometrioid tumors), germ cell (e.g., teratomas, choriocarcinomas, polyembryomas, embryomal carcinoma, endodermal sinus tumor, dysgerminoma, gonadoblastoma), stromal carcinomas (e.g., granulosal stromal cell tumors)), fallopian tube carcinoma, peritoneal carcinoma, leiomyoma
pancreas:	adenocarcinoma, ductal adenocarcinoma, mucinous cyst adenocarcinoma, acinar cell carcinoma, unclassified large cell carcinoma, small cell carcinoma, pancreatoblastoma, duct-ectatic mucin-hypersecreting tumor, mucinous cyst adenoma, papillary cystic neoplasm, serous cyst adenoma, diabetes melitis, chronic pancreatitis
prostate:	epithelial neoplasms (e.g., adenocarcinoma, small cell tumors, transitional cell carcinoma, carcinoma in situ, and basal cell carcinoma), carcinosarcoma, non-epithelial neoplasms (e.g., mesenchymal and lymphoma), germ cell tumors, prostatic intraepithelial neoplasia (PIN), hormone independent prostate cancer, benign prostate hyperplasia, prostatitis
skin/melanoma:	melanoma, lentigo (common benign localized hyperplasia of melanocytes), nevocellular nevi (congenital or acquired neoplasm of melanocytes), actinic keratosis (overgrowth of outer layers of skin), basal cell carcinoma, Merkel cell carcinoma, benign fibrous histiocytoma (dermal neoplasms of fibroblasts and histiocytes), dermatofibrosarcoma protuberans (well differentiated fibrosarcoma of the skin), xanthomas (tumor-like collections of foamy histiocytes within the dermis), dermal vascular tumors, seborrheic keratoses (benign tumor), acanthosis nigricans (benign or malignant hyperplasia and hyperpigmentation of skin), and squamous cell carcinomas of the skin, lung, cervix, esophagus, uterus, head, neck, or bladder
soft tissue:	soft tissue tumors (e.g., fibrosarcoma, liposarcoma, leiomyosarcoma, histiocytoma, fibrohistiocytic sarcoma) smooth muscle tumors (e.g., rhabdomyoma, rhabdomyosarcoma) tumors of the blood and lymph vessels (e.g., angiosarcoma, lymphangiosarcoma, Kaposi's sarcoma), perivascular tumors (e.g., glomus tumors, hemangiopericytoma), synovial tumors (e.g., mesothelioma), neural tumors (e.g., neurofibroma, neurofibrosarcoma, malignant peripheral nerve sheath tumors, granular cell tumors, plexosarcoma, ganglioneuroblastoma, neuroepithelioma, extraskeletal Ewing's sarcoma, schwannoma, neuroma, ganglioneuroma), paraganglioma, extraskeletal cartilaginous and osseous tumors (e.g., chondrosarcoma, osteosarcoma), pluripotential mesenchymal tumors, epitheliod sarcomas, rhabdoid tumors, desmoplastic small cell tumors, alveolar sarcoma
stomach:	adenocarcinoma, squamous cell carcinoma, adenoacanthoma, carcinoid, leiomyosarcoma, gastritis (chronic atrophic, H. pylori associated), hyperplastic polyps, lipoma, leiomyoma, esophageal adenocarcinomas
testicles:	germ cell tumors (including seminomas, embryonal carcinomas, teratomas, choriocarcinomas, yolk sac tumors), sex chord stromal tumors (including Leydig cell tumors, Sertoli cell tumors, and Granulosa cell tumors), germ cell and gonadal stromal elements (e.g., gonadoblastomas), adnexal and paratesticular tumors (e.g., mesotheliomas, soft tissue sarcomas, and adnexal of the rete testes), miscellaneous neoplasms (including carcinoid, lymphoma, and cysts)
uterus:	epithelial tumors (e.g., endometrioid, papillary endometrioid, papillary serous, clear cell, mucinous), mesenchymal tumors (e.g., endometrial stromal sarcoma, leiomyosarcoma, nonspecific sarcomas), mixed tumors (e.g., malignant mixed mullerian tumors, adenosarcoma)

Table 2: Disease Indications of Selected Genes

Table 2 provides disease indications for about 59 selected genes. These genes may be useful as targets for small molecule, antibody, or DNA vaccine therapy. They may also have utility as prognostic or diagnostic markers. These genes were identified using Eos/Affymetrix Genechip arrays. The columns in Table 2 are as follows:

Pkey: Unique Eos probeset identifier number

Ex Accn: Exemplar Accession number

UnigeneID: UniGene ID number

UnigeneTitle: UniGene title

Disease Indications: Diseases indicated for selected gene as described in Table 1 and abbreviated as follows:

AWPC (androgen independent prostate diseases), arth (arthritic diseases), bph (benign prostatic hyperplasia), blad (bladder diseases), angio (blood vessel diseases), EWS (bone diseases), glio (brain diseases), breast (breast diseases), cerv (cervical diseases), colon (colorectal diseases), esoph (esophageal diseases), fibro (fibrotic diseases), headnk (head & neck diseases), leio (leiomyoma diseases), leuk (leukocyte diseases), hepC (liver diseases), lung (lung diseases), ovar (ovarian diseases), endo (ovarian endometrioid diseases), omuc (ovarian mucinous diseases), panc (pancreatic diseases), pros (prostate diseases), renal (renal diseases), mela (skin diseases), stom (stomach diseases), test (testicular diseases), uter (uterine diseases)

AA: Refseq amino acid accession number

NA: Refseq nucleotide accession number

SEQ ID NOs: Sequence identification numbers linking Pkey to corresponding SEQ ID NOs:1-116.

Table 2: Disease Indications of Selected Genes

Pkey	Ex Accn	UnigeneID	Unigene Title	Disease Indications	NA	AA	SEQ ID NOs.
453983	H94997	Hs.318751	ESTs	angio	FGENESH	FGENESH	Seq ID No. 1 & 59
453983	H94997	Hs.318751	ESTs	angio	NM_020249.1	NP_064634.1	Seq ID No. 2 & 60
428758	AA433988	Hs.98502	CA125 antigen; mucin 16	ovar, cerv, lung, panc, stom, renal	NM_002253.1	NP_002244.1	Seq ID No. 3 & 61
450983	AA305384	Hs.25740	ERO1 (S. cerevisiae)-like	blad, lung, ovar, panc	NM_014584.1	NP_055399.1	Seq ID No. 4 & 62
417771	AA804698	Hs.82547	retinoic acid receptor responder (tazaro	blad, cerv, panc, pros, ovar	NM_002888.1	NP_002879.1	Seq ID No. 5 & 63
448262	AW880830	Hs.186273	Homo sapiens quiescin Q6 (QSCN6)	blad	NM_002826.2	NP_002817.2	Seq ID No. 6 & 64
407720	AB037776	Hs.38002	immunoglobulin superfamily, member 9	lung	NM_020789.1	NP_065840.1	Seq ID No. 7 & 65
435013	H91923	Hs.110024	NM_020142:Homo sapiens NADH:ubiquinone o	renal, lung, sarc	NM_020142.2	NP_064527.1	Seq ID No. 8 & 66
330844	AA063037	Hs.66803	ESTs	lung	NM_016247.1	NP_057331.1	Seq ID No. 9 & 67
440659	AF134160	Hs.7327	claudin 1	lung	NM_021101	NP_066924.1	Seq ID No. 10 & 68

449101	AA205847	Hs.23016	G protein-coupled receptor	lung, headnk	XM_051522.4	XP_051522.2	Seq ID No. 11 & 69
429263	AA019004	Hs.198396	ATP-binding cassette, sub-family A (ABC1	lung	NM_000350.1	NP_000341.1	Seq ID No. 12 & 70
421474	U76362	Hs.104637	solute carrier family 1 (glutamate trans	lung	NM_006671.2	NP_006662.2	Seq ID No. 13 & 71
421753	BE314828	Hs.107911	ATP-binding cassette, sub-family B (MDR/	lung	NM_005689	NP_005680.1	Seq ID No. 14 & 72
408482	NM_000676	Hs.45743	adenosine A2b receptor	lung, esoph, headnk, colon	NM_000676	NP_000667.1	Seq ID No. 15 & 73
426761	AI015709	Hs.172089	PORIMIN Pro-oncosis receptor inducing me	lung, esoph, pros, uter, panc, colon, ovar, headnk	NM_052932	NP_443164	Seq ID No. 16 & 74
429736	AF125304	Hs.212680	tumor necrosis factor receptor superfami	lung	NM_004195	NP_004186.1	Seq ID No. 17 & 75
430985	AA490232	Hs.27323	ESTs, Weakly similar to I78885 serine/th	lung	AK091896.1	BAC03767.1	Seq ID No. 18 & 76
431890	X17033	Hs.271986	integrin, alpha 2 (CD49B, alpha 2 subuni	blad, headnk, lung, panc, cerv, stom	NM_002203.2	NP_002194.1	Seq ID No. 19 & 77
432583	AW023624	Hs.162282	potassium channel TASK-4; potassium chan	lung	NM_031460	NP_113648.1	Seq ID No. 20 & 78
446872	X97058	Hs.16362	pyrimidinergic receptor P2Y, G-protein c	lung	NM_004154	NP_004145.1	Seq ID No. 21 & 79
453102	NM_007197	Hs.31664	frizzled (Drosophila) homolog 10	lung, headnk, colon	NM_007197	NP_009128.1	Seq ID No. 22 & 80
404287	NM_173674.1	Hs.449321	Homo sapiens discoidin, CUB and LCCL domain containing 1 (DCBLD1)	panc, lung, colon, uter, esoph	NM_173674.1	NP_775945.1	Seq ID No. 23 & 81
404287	NM_173674.1	Hs.449321	Homo sapiens discoidin, CUB and LCCL domain containing 1 (DCBLD1)	panc, lung, colon, uter, esoph	NM_173674.1	NP_775945.1	Seq ID No. 24 & 82
418318	U47732	Hs.84072	transmembrane 4 superfamily member 3	panc, pros, colon, stom, omuc	NM_004616.2	NP_004607.1	Seq ID No. 25 & 83
444754	T83911	Hs.11881	transmembrane 4 superfamily member 4	panc, omuc, stom, lung, colon	NM_004617.2	NP_004608.1	Seq ID No. 26 & 84
428505	AL035461	Hs.2281	chromogranin B (secretogranin 1)	panc, lung	NM_001819	NP_001810.1	Seq ID No. 27 & 85
448844	AI581519	Hs.177164	FGENESH predicted novel cell surface_pr	panc, lung, stom, omuc	XM_093082.1	XP_093082.1	Seq ID No. 28 & 86
448844	AI581519	Hs.177164	FGENESH predicted novel cell surface_pr	panc, lung, stom, omuc	FGENESH	FGENESH	Seq ID No. 29 & 87
426227	U67058	Hs.154299	Human proteinase activated receptor-2 mR	panc, lung, colon, esoph, stom	NM_005242.2	NP_005233.2	Seq ID No. 30 & 88
445417	AK001058	Hs.12680	a disintegrin-like and metalloprotease w	panc, headnk, stom, lung, esoph, sarc, colon	NM_030955	NP_112217.1	Seq ID No. 31 & 89
413719	BE439580	Hs.75498	small inducible cytokine subfamily A (Cy	leuk, panc, lung, headnk, cerv, colon, uter, stom, esoph	NM_004591	NP_004582.1	Seq ID No. 32 & 90
416498	U33632	Hs.79351	potassium channel, subfamily K, member 1	panc, stom, breast, endo, colon	NM_002245.2	NP_002236.1	Seq ID No. 33 & 91
413095	AA494359	Hs.30715	potassium voltage-gated channel, Isk-rel	panc, stom, renal, colon	NM_005472.1	NP_005463.1	Seq ID No. 34 & 92

426125	X87241	Hs.166994	FAT tumor suppressor (Drosophila) homolo	colon, stom, panc, pros, renal, fibro, cerv	NM_005245.1	NP_005236.1	Seq ID No. 35 & 93
436729	BE621807	Hs.351316	transmembrane 4 superfamily member 1	panc, colon, stom, ovar, lung, blad	NM_014220.1	NP_055035.1	Seq ID No. 36 & 94
437145	AF007216	Hs.5462	solute carrier family 4, sodium bicarbon	panc, pros, stom	NM_003759.1	NP_003750.1	Seq ID No. 37 & 95
451820	AW058357	Hs.199248	ESTs	panc	NM_000958	NP_000949.1	Seq ID No. 38 & 96
427557	NM_002659	Hs.179657	plasminogen activator, urokinase recepto	panc, colon, stom, ovar, cerv, blad, lung, headnk, esoph	NM_002659.1	NP_002650.1	Seq ID No. 39 & 97
408308	AL033377	Hs.44197	hypothetical protein DKFZp564D0462	panc, renal, colon	AK027843.1	BAB55406.1	Seq ID No. 40 & 98
428242	H55709	Hs.2250	leukemia inhibitory factor (cholinergic	ovar, panc, leuk, lung	NM_002309.2	NP_002300.1	Seq ID No. 41 & 99
428778	AK000530	Hs.193326	fibroblast growth factor receptor-like 1	ovar	NM_021923	NP_068742	Seq ID No. 42 & 100
439659	AW970780	Hs.59483	leucine-rich repeat-containing G protein	ovar, stom, mela, colon	XM_097508	XP_097508	Seq ID No. 43 & 101
411825	AK000334	Hs.352415	solute carrier family 39 (zinc transport	colon, ovar	NM_130849	NP_570901	Seq ID No. 44 & 102
442133	AW874138	Hs.129017	ESTs; type Ia transmembrane protein	ovar, uter	XM_087172	XP_087172	Seq ID No. 45 & 103
412314	AA825247	Hs.356084	G protein-coupled receptor 27 (GPR27) (S	ovar, uter, test	NM_018971	NP_061844	Seq ID No. 46 & 104
411828	AW161449	Hs.72290	wingless-type MMTV integration site fami	ovar	NM_004625	NP_004616	Seq ID No. 47 & 105
439668	AI091277	Hs.302634	frizzled (Drosophila) homolog 8	ovar, uter	NM_031866	NP_114072	Seq ID No. 48 & 106
433336	AF017986	Hs.31386	secreted frizzled-related protein 2 (str	ovar, fibro, headnk, lung, panc, blad	XM_050625	XP_050625	Seq ID No. 49 & 107
432128	AA127221	Hs.66	Interleukin 1 receptor-like 1	angio	BC030975.1	AAH30975.1	Seq ID No. 50 & 108
446921	AB012113	Hs.16530	small inducible cytokine subfamily A (Cy	breast, panc, headnk, lung, fibro, mela	NM_002988.1	NP_002979.1	Seq ID No. 51 & 109
450623	H02562	Hs.28848	Nedd4 binding protein 3 (N4BP3)	angio	XM_038920.3	XP_038920.2	Seq ID No. 52 & 110
450623	H02562	Hs.28848	Nedd4 binding protein 3 (N4BP3)	angio	FGENESH	FGENESH	Seq ID No. 53 & 111
432179	X75208	Hs.2913	EphB3	ovar, colon, lung, pros	NM_004443	NP_004434.1	Seq ID No. 54 & 112
431870	AW449902	Hs.105500	Homo sapiens POU domain, class 5, transc	renal	FGENESH	FGENESH	Seq ID No. 55 & 113
431870	AW449902	Hs.105500	Homo sapiens POU domain, class 5, transc	renal	XM_175178.1	XP_175178.1	Seq ID No. 56 & 114
437212	AI765021	Hs.210775	ESTs	renal, uter, ovar	NM_001074.1	NP_001065.1	Seq ID No. 57 & 115
442438	AA995998	Hs.371863	gb:os26b03.s1 NCI_CGAP_Kid5 Homo sapiens	uter, ovar,renal	FGENESH	FGENESH	Seq ID No. 58 & 116

It is understood that the examples described above in no way serve to limit the true scope of this invention, but rather are presented for illustrative purposes. All publications, sequences of accession numbers, and patent applications cited in this specification are herein incorporated by reference as if each individual publication, accession number, or

patent application were specifically and individually indicated to be incorporated by reference.

CLAIMS

WHAT IS CLAIMED IS:

1. A method for detecting a pathological cell in a patient, said method comprising detecting in a biological sample from said patient a nucleic acid or polypeptide comprising a sequence at least 80% identical to a sequence selected from SEQ ID NOs:1-116.

2. The method of Claim 1, wherein said pathological cell has a pathology selected from those listed Table 1.

3. The method of Claim 1, wherein said biological sample is tissue from an organ which is affected by a pathology listed in Table 1.

4. The method of Claim 1, wherein said nucleic acids are mRNA.

5. The method of Claim 1, further comprising a step of amplifying nucleic acids.

6. The method of Claim 1, wherein said nucleic acid comprises a sequence selected from SEQ ID NOs:1-58.

7. The method of Claim 1, wherein said polypeptide comprises a sequence selected from SEQ ID NOs:59-116.

8. The method of Claim 1, wherein said detecting comprises using a biochip comprising a nucleic acid at least 80% identical to SEQ ID NOs:1-58.

9. The method of Claim 1, wherein said patient is undergoing a therapeutic regimen to treat a pathology selected from those listed Table 1.

10. The method of Claim 1, wherein said patient is suspected of having a pathology selected from those listed Table 1.

11. An isolated nucleic acid molecule comprising a sequence selected from SEQ ID NOs: 1-58.

12. The nucleic acid molecule of Claim 11, wherein the nucleic acid is labeled.

13. An expression vector comprising the nucleic acid of Claim 11.

14. A host cell comprising the expression vector of Claim 13.

15. An isolated nucleic acid encoding a polypeptide sequence selected from SEQ ID NOs: 59-116.

16. An isolated polypeptide encoded by a sequence selected from SEQ ID NOs:1-58.

17. An antibody that specifically binds a polypeptide of Claim 16.

18. The antibody of Claim 17, wherein the antibody is a humanized antibody.

19. The antibody of Claim 17, wherein the antibody is an antibody fragment.

20. The antibody of Claim 17, wherein the antibody is conjugated to an effector component.

21. The antibody of Claim 17, wherein the antibody is conjugated to a detectable label or a cytotoxic chemical.

22. A method for specifically targeting a compound to a pathological cell in a patient, said method comprising administering to said patient an antibody of Claim 17, wherein said antibody is conjugated to the compound.

23. A method for detecting a pathological cell in a patient, said method comprising contacting a biological sample with an antibody of Claim 17.

24. The method of Claim 22, wherein said antibody is conjugated to an effector component or a fluorescent label.

25. The method of Claim 22, wherein said said biological sample is a blood, serum, urine, or stool sample.

26. A method for identifying a compound that modulates a pathology-associated polypeptide, said method comprising:

a) contacting said compound with a pathology-associated polypeptide, said polypeptide encoded by a polynucleotide that selectively hybridizes to a sequence at least 80% identical to SEQ ID NOs:1-58; and

b) determining the effect of said compound upon the function of said polypeptide.

27. A screening assay comprising:

a) administering a test compound to a cell from a mammal exhibiting a pathology selected from those listed in Table 1;

b) administering a test compound to a cell from a mammal not exhibiting said pathology;

c) comparing the expression level of a polynucleotide of the cell comprising a sequence at least 80% identical to SEQ ID NOs:1-58 with the expression level of said polynucleotide of a control cell;

whereby modulation of the expression level of the polynucleotide of the cell indicates that the test compound is a drug candidate.

SEQUENCE LISTING

<110> Aziz, Natasha
Gish, Kurt
Wilson, Keith
Zlotnik, Albert

<120> METHODS OF DIAGNOSIS OF CANCER, COMPOSITIONS AND

<130> 05882.0191.00PC00

<160> 116

<170> PatentIn version 3.2

<210> 1
<211> 7008
<212> DNA
<213> Homo Sapiens

<400> 1
atggctatga tgattcttcg ggttgactac acatttgagg aaaatagaga caagttagct 60

tccaggaaga aggaatacag tcaaggaagt gtggcagacc tgactccaga caattggaaa 120

aacatcaccg tgcctcacag tggaagacat tcagaggtgt ctaggggaga gctggtctgc 180

agaacttgct cagaatgttc agctggtccc cacatctgga tgaaagggct ctatcagacc 240

caagatgaag aagcaggagg agaaaatatt ttcattctgt tgttcattga gtcaacacaa 300

tttggacagt ttgtggccat gggctctccg atcacagaac ataaagtctt taccatgtat 360

cttggtttag ccacacatct attttacagc cttataactc acccttttgt tcttttggaa 420

aaccactcct gcccaagctc agtccatggg tttgatgtag ctgggctgat ctttgacaaa 480

gtgggcatga gatccagacc tggccggatg ggagcactgt ttgcatattt tgctggattt 540

atcaggagaa aggcactggt tgtttgtttg tttgtttttt gctggagtaa tgaagctgct 600

aacaagcccc ccattcaaga agccgctcag ctatcccggc cagcacaggg cgcccggcgc 660

gcctcggagc gcaagttcct cgccttctcc tgcccgctcg ctgggcatta tgcggccaag 720

cagccgagcc ccagtcctcc tcctcctcct gctcctccgg ctcctcctgc ggcccgagcg 780

gctcagctct cggcaggcgg cggcgttgct cagccgagcg cagacgggac cctcgcagcg 840

agacctcagc gactcctaaa gtcaaaagtt ggcggcgggc gccgggctcc gcgcgctctc 900

cacggccgct gcctcgcgtc gccgccgcag ccaaggaggg caggagggag gggggtgggg 960

gcagcggagg gaggggtggg aagcaccatg cagtttgtat cctgggccac actgctaacg 1020

ctcctggtgc gggacctggc cgagatgggg agcccagacg ccgcggcggc cgtgcgcaag 1080

gacaggctgc acccgaggca agtgaaatta ttagagaccc tgagcgaata cgaaatcgtg 1140

tctcccatcc gagtgaacgc tctcggagaa ccctttccca cgaacgtcca cttcaaaaga 1200

acgcgacgga gcattaactc tgccactgac ccctggcctg ccttcgcctc ctcctcttcc 1260

tcctctacct cctcccaggc gcattaccgc ctctctgcct tcggccagca gtttctattt 1320

aatctcaccg ccaatgccgg atttatcgct ccactgttca ctgtcaccct cctcgggacg 1380

cccggggtga atcagaccaa gttttattcc gaagaggaag cggaactcaa gcactgtttc 1440

tacaaaggct atgtcaatac caactccgag cacacggccg tcatcagcct ctgctcagga 1500

atggggcttc tggatgtttc agagctttct ggagtttgga ctcggttcag cggcgcgttg 1560

cccaacgctg cgaggcggcc tggaagtcag tttccaaact cggagaaagt tactggagtt 1620

gcggttccct gcagcaaact tggccacccg ggagcggagc cgctgagcgc agggcggacc 1680

agactcctta ttgtggatct tacaaggcat ctgcccccca cttctccacg acatttgagg 1740

agtcgctgcg gtacggttct ggcccgagca agggtggtcc ttgactttcc caagcgacgt 1800

gcgtttctgc cacgcgcgtg tgacgcagaa actttcccgg cggggccttg gatactcacc 1860

cccagacact gggctgctcc atcagtgcgc tgtcgttcgt gggtcttaaa gttcccaagc 1920

acatccttcc ttctctgcct ttcaatggaa ggatcgggtg gagagcgtgg caagcctgaa 1980

gactgggagg gggtagttct agcctgctgg gattcaagga aagggataaa ccccttagc 2040

ccccagcaaa gcgcccggag ccgtggctcc cgaaatgcgc tgtcgagatt gtttgggggt 2100

gggcggagac ggcagcttgg cgaagtggga gggggtgctg cactgggcac attccggtct 2160

catgatgggg attattttat tgaaccacta cagtctatgg atgaacaaga agatgaagag 2220

gaacaaaaca aaccccacat catttatagg cgcagcgccc cccagagaga gccctcaaca 2280

ggaaggcatg catgtgacac ctcagggtta cagaaatgtc ttataaatgg aagccacgaa 2340

aacatatatg tgtttgttga atgtttccta gaaacttcag gtttgctcat gttctgtgac 2400

ttaaggaact gtagcaaggt acctgtacgt tatgctgtga gctacttctg cacccccttct 2460

ttgaattcag atgcagcttc tcagaacagt ttagaatatg gcacgattca ccagcaggta 2520

tcagaggaat ggaccaacag gtcaaggaca cctctggaac cagaacacaa aaataggcac 2580

agtaaagaca agaagaaaac cagagcaaga aaatggggag aaaggattaa cctggctggt 2640

gacgtagcag cattaaacag cggcttagca acagaggcat tttctgctta tggtaataag 2700

acggacaaca caagagaaaa gaggacccac agaaggacaa aacgtttttt atcctatcca 2760

cggtttgtag aagtcttggt ggtggcagac aacagaatgg tttcatacca tggagaaaac 2820

cttcaacact atattttaac tttaatgtca attgtagcct ctatctataa agacccaagt 2880

attggaaatt taattaatat tgttattgtg aacttaattg tgattcataa tgaacaggat 2940

gggccttcca tatcttttaa tgctcagaca acattaaaaa acttttgcca gtggcagcat 3000

tcgaagaaca gtccaggtgg aatccatcat gatactgctg ttctcttaac aagacaggat 3060

atctgcagag ctcacgacaa atgtgatacc ttaggcctgg ctgaactggg aaccatttgt 3120

gatccctata gaagctgttc tattagtgaa gatagtggat tgagtacagc ttttacgatc 3180

gcccatgagc tgggccatgt gtttaacatg cctcatgatg acaacaacaa atgtaaagaa 3240

gaaggagtta agagtcccca gcatgtcatg gctccaacac tgaacttcta caccaacccc 3300

tggatgtggt caaagtgtag tcgaaaatat atcactgagt ttttagacac tggttatggc 3360

gagtgtttgc ttaacgaacc tgaatccaga ccctaccctt tgcctgtcca actgccaggc 3420

atcctttaca acgtgaataa acaatgtgaa ttgatttttg gaccaggttc tcaggtgtgc 3480

ccatatatgc actgcaagta tggattttgt gttcccaaag aaatggatgt ccccgtgaca 3540

gatggatcct ggggaagttg gagtcccttt ggaacctgct ccagaacatg tggaggggggc 3600

atcaaaacag ccattcgaga gtgcaacaga ccagaaccaa aaaatggtgg aaaatactgt 3660

gtaggacgta gaatgaaatt taagtcctgc aacacggagc catgtctcaa gcagaagcga 3720

gacttccgag atgaacagtg tgctcacttt gacgggaagc attttaacat caacggtctg 3780

cttcccaatg tgcgctgggt ccctaaatac agtggaattc tgatgaagga ccggtgcaag 3840

ttgttctgca gagtggcagg gaacacagcc tactatcagc ttcgagacag agtgatagat 3900

ggaactcctt gtggccagga cacaaatgat atctgtgtcc agggcctttg ccggcaagct 3960

ggatgcgatc atgttttaaa ctcaaaagcc cggagagata aatgtggggt ttgtggtggc 4020

gataattctt catgcaaaac agtggcagga acatttaata cagtacatta tggttacaat 4080

actgtggtcc gaattccagc tggtgctacc aatattgatg tgcggcagca cagtttctca 4140

ggggaaacag acgatgacaa ctacttagct ttatcaagca gtaaaggtga attcttgcta 4200

aatggaaact ttgttgtcac aatggccaaa agggaaattc gcattgggaa tgctgtggta 4260

gagtacagtg ggtccgagac tgccgtagaa agaattaact caacagatcg cattgagcaa 4320

gaacttttgc ttcaggtttt gtcggtggga aagttgtaca accccgatgt acgctattct 4380

ttcaatattc caattgaaga taaacctcag cagttttact ggaacagtca tgggccatgg 4440

caagcatgca gtaaaccctg ccaaggggaa cggaaacgaa aacttgtttg caccagggaa 4500

tctgatcagc ttactgtttc tgatcaaaga tgcgatcggc tgccccagcc tggacacatt 4560

actgaaccct gtggtacaga ctgtgacctg aggtggcatg ttgccagcag gagtgaatgt 4620

agtgcccagt gtggcttggg ttaccgcaca ttggacatct actgtgccaa atatagcagg 4680

ctggatggga agactgagaa ggttgatgat ggtttttgca gcagccatcc caaaccaagc 4740

aaccgtgaaa aatgctcagg ggaatgtaac acgggtggct ggcgctattc tgcctggact 4800

gaatgttcaa aaagctgtga cggtgggacc cagaggagaa gggctatttg tgtcaatacc 4860

cgaaatgatg tactggatga cagcaaatgc acacatcaag agaaagttac cattcagagg 4920
tgcagtgagt tcccttgtcc acagtggaaa tctggagact ggtcagagtg cttggtcacc 4980
tgtggaaaag ggcataagca ccgccaggtc tggtgtcagt ttggtgaaga tcgattaaat 5040
gatagaatgt gtgaccctga gaccaagcca acatctatgc agacttgtca gcagccggaa 5100
tgtgcatcct ggcaggcggg tccctgggga cagtgcagtg tcacttgtgg acagggatac 5160
cagctaagag cagtgaaatg catcattggg acttatatgt cagtggtaga tgacaatgac 5220
tgtaatgcag caactagacc aactgatacc caggactgtg aattaccatc atgtcatcct 5280
cccccagctg ccccggaaac gaggagaagc acatacagtg caccaagaac ccagtggcga 5340
tttgggtctt ggaccccatg ctcagccact tgtgggaaag gtacccggat gagatacgtc 5400
agctgccgag atgagaatgg ctctgtggct gacgagagtg cctgtgctac cctgcctaga 5460
ccagtggcaa aggaagaatg ttctgtgaca ccctgtgggc aatggaaggc cttggactgg 5520
agctcttgct ctgtgacctg tgggcaaggt agggcaaccc ggcaagtgat gtgtgtcaac 5580
tacagtgacc acgtgatcga tcggagtgag tgtgaccagg attatatccc agaaactgac 5640
caggactgtt ccatgtcacc atgccctcaa aggaccccag acagtggctt agctcagcac 5700
cccttccaaa atgaggacta tcgtccccgg agcgccagcc ccagccgcac ccatgtgctc 5760
ggtggaaacc agtggagaac tggcccctgg ggagcaacat attggagaga gaataccatg 5820
gagtttctag agctgtttct tccagaatct ttaactggac caggtagcaa atcctgtgac 5880
cagcactatg gaagtacctg tgctggcgga tcccagcggc gtgttgttgt atgtcaggat 5940
gaaaatggat acaccgcaaa cgactgtgtg gagagaataa aacctgatga gcaaagagcc 6000
tgtgaatccg gcccttgtcc tcagtgggct tatggcaact ggggagagtg cactaagctg 6060
tgtggtggag gcataagaac aagactggtg gtctgtcagc ggtccaacgg tgaacggttt 6120
ccagatttga gctgtgaaat tcttgataaa cctcccgatc gtgagcagtg taacacacat 6180
gcttgtccac acgacgctgc atggagtact ggcccttgga gctcgtccat gtggcaggtg 6240
aataataaaa cagttacgct tgggaacttg tgttctgtct cttgtggtcg agggcataaa 6300
caacgaaatg tttactgcat ggcaaaagat ggaagccatt tagaaagtga ttactgtaag 6360
cacctggcta agccacatgg gcacagaaag tgccgaggag gaagatgccc caaatggaaa 6420
gctggcgctt ggagtcagaa aactactaac tcagactgca ctgaagctga ctgtggtcac 6480
ctggcagaaa ttgagtctca gtttatcttg gaggttcttg aagaaagggc tgttgacgaa 6540
agttctagaa aatacctctg cccatttgct tgcttacaaa agtgctctgt gtcctgtggc 6600
cgaggcgtac agcagaggca tgtgggctgt cagatcggaa cacacaaaat agccagagag 6660

accgagtgca acccatacac cagaccggag tcggaacgcg actgccaagg cccacggtgt 6720

cccctctaca cttggagggc agaggaatgg caagaaacct accatggcct gctctctcca 6780

tctccctctt tgtgtcacgc taaactcaac cctgctccga ggagcggaaa gcctcaacct 6840

agatgtcact tcctctcaga agcctttgcc aatcacacca ccccactaaa tctgagtcag 6900

atgctcctcc actcagctct cacaacacac gcagattact gtactctggc agttaacacc 6960

tggaattctc attgcctgtt tttctcatct atgttatcag ttatttaa 7008

<210> 2
<211> 3674
<212> DNA
<213> Homo Sapiens

<400> 2
gcgggaagca ccatgcagtt tgtatcctgg gccacactgc taacgctcct ggtgcgggac 60

ctggccgaga tggggagccc agacgccgcg gcggccgtac gcaaggacag gctgcacccg 120

aggcaagtga aattattaga gaccctgggc gaatacgaaa tcgtgtctcc catccgagtg 180

aacgctctcg gagaaccctt tcccacgaac gtccacttca aaagaacgcg acggagcatt 240

aactctgcca ctgacccctg gcctgccttc gcctcctcct cttcctcctc tacctcctcc 300

caggcgcatt accgcctctc tgccttcggc cagcagtttc tatttaatct caccgccaat 360

gccggattta tcgctccact gttcactgtc accctcctcg ggacgcccgg ggtgaatcag 420

accaagtttt attccgaaga ggaagcggaa ctcaagcact gtttctacaa aggctatgtc 480

aataccaact ccgagcacac ggccgtcatc agcctctgct caggaatgct gggcacattc 540

cggtctcatg atggggatta ttttattgaa ccactacagt ctatggatga acaagaagat 600

gaagaggaac aaaacaaacc ccacatcatt tataggcgca gcgccccca gagagagccc 660

tcaacaggaa ggcatgcatg tgacacctca gaacacaaaa ataggcacag taaagacaag 720

aagaaaacca gagcaagaaa atggggagaa aggattaacc tggctggtga cgtagcagca 780

ttaaacagcg gcttagcaac agaggcattt tctgcttatg gtaataagac ggacaacaca 840

agagaaaaga ggacccacag aaggacaaaa cgtttttat cctatccacg gtttgtagaa 900

gtcttggtgg tggcagacaa cagaatggtt tcataccatg gagaaaacct tcaacactat 960

attttaactt taatgtcaat tgtagcctct atctataaag acccaagtat tggaaattta 1020

attaatattg ttattgtgaa cttaattgtg attcataatg aacaggatgg gccttccata 1080

tcttttaatg ctcagacaac attaaaaaac ctttgccagt ggcagcattc gaagaacagt 1140

ccaggtggaa tccatcatga tactgctgtt ctcttaacaa gacaggatat ctgcagagct 1200

cacgacaaat gtgatacctt aggcctggct gaactgggaa ccatttgtga tccctataga 1260

agctgttcta ttagtgaaga tagtggattg agtacagctt ttacgatcgc ccatgagctg 1320

ggccatgtgt ttaacatgcc tcatgatgac aacaacaaat gtaaagaaga aggagttaag 1380

agtccccagc atgtcatggc tccaacactg aacttctaca ccaacccctg gatgtggtca 1440

aagtgtagtc gaaaatatat cactgagttt ttagacactg gttatggcga gtgtttgctt 1500

aacgaacctg aatccagacc ctacccttta cctgtccaac tgccaggcat cctttacaac 1560

gtgaataaac aatgtgaatt gatttttgga ccaggttctc aggtgtgccc atatatgatg 1620

cagtgcagac ggctctggtg caataacgtc aatggagtac acaaaggctg ccggactcag 1680

cacacaccct gggccgatgg gacggagtgc gagcctggaa agcactgcaa gtatggattt 1740

tgtgttccca aagaaatgga tgtccccgtg acagatggat cctggggaag ttggagtccc 1800

tttggaacct gctccagaac atgtggaggg ggcatcaaaa cagccattcg agagtgcaac 1860

agaccagaac caaaaaatgg tggaaaatac tgtgtaggac gtagaatgaa atttaagtcc 1920

tgcaacacgg agccatgtct caagcagaag cgagacttcc gagatgaaca gtgtgctcac 1980

tttgacggga agcattttaa catcaacggt ctgcttccca atgtgcgctg ggtccctaaa 2040

tacagtggaa ttctgatgaa ggaccggtgc aagttgttct gcagagtggc agggaacaca 2100

gcctactatc agcttcgaga cagagtgata gatggaactc cttgtggcca ggacacaaat 2160

gatatctgtg tccagggcct ttgccggcaa gctggatgcg atcatgtttt aaactcaaaa 2220

gcccggagag ataaatgtgg ggtttgtggt ggcgataatt cttcatgcaa aacagtggca 2280

ggaacattta atacagtaca ttatggttac aatactgtgg tccgaattcc agctggtgct 2340

accaatattg atgtgcggca gcacagtttc tcaggggaaa cagacgatga caactactta 2400

gctttatcaa gcagtaaagg tgaattcttg ctaaatggaa actttgttgt cacaatggcc 2460

aaaagggaaa ttcgcattgg gaatgctgtg gtagagtaca gtgggtccga gactgccgta 2520

gaaagaatta actcaacaga tcgcattgag caagaacttt tgcttcaggt tttgtcggtg 2580

ggaaagttgt acaaccccga tgtacgctat tctttcaata ttccaattga agataaacct 2640

cagcagtttt actggaacag tcatgggcca tggcaagcat gcagtaaacc ctgccaaggg 2700

gaacggaaac gaaaacttgt ttgcaccagg gaatctgatc agcttactgt ttctgatcaa 2760

agatgcgatc ggctgcccca gcctggacac attactgaac cctgtggtac agactgtgac 2820

ctgaggtggc atgttgccag caggagtgaa tgtagtgccc agtgtggctt gggttaccgc 2880

acattggaca tctactgtgc caaatatagc aggctggatg ggaagactga gaaggttgat 2940

gatggttttt gcagcagcca tcccaaacca agcaaccgtg aaaaatgctc aggggaatgt 3000

aacacgggtg gctggcgcta ttctgcctgg actgaatgtt caaaaagctg tgacggtggg 3060

acccagagga gaagggctat ttgtgtcaat acccgaaatg atgtactgga tgacagcaaa 3120

tgcacacatc aagagaaagt taccattcag aggtgcagtg agttcccttg tccacagtgg 3180

aaatctggag actggtcaga ggtaagatgg gagggctgtt atttccccta ggtcatctct 3240

tacattctag ttctggtgct ctctatctgt ttaagacaaa cccttgtgca cctttctccc 3300

acctctccct ttctcccttg tctcccttga gaaaacaact ccagttctct gcctgcacca 3360

tgactgtcgt actggatgta actagtctac cagtgacctc agggcacttt gggcttggct 3420

agatcactca ctgttgtagc ttctgttgtg attttgaagt tgcagtccat caccttccct 3480

cctctttgag ccctagctaa gtcactgaaa ggaaatcatg gatttattaa tcataaagct 3540

atactagctc acatctgaag tcaacatgaa gtttcctact tccttgtctt tgaaataaga 3600

gaattagacc ccagggagtg acctctctga cttacccatc caactgccca aaaaaaaaaa 3660

aaaaaaaaaa aaaa 3674

<210> 3
<211> 5830
<212> DNA
<213> Homo Sapiens

<400> 3
actgagtccc gggacccccgg gagagcggtc agtgtgtggt cgctgcgttt cctctgcctg 60

cgccgggcat cacttgcgcg ccgcagaaag tccgtctggc agcctggata tcctctccta 120

ccggcacccg cagacgcccc tgcagccgcc ggtcggcgcc cgggctccct agccctgtgc 180

gctcaactgt cctgcgctgc ggggtgccgc gagttccacc tccgcgcctc cttctctaga 240

caggcgctgg gagaaagaac cggctcccga gttctgggca tttcgcccgg ctcgaggtgc 300

aggatgcaga gcaaggtgct gctggccgtc gccctgtggc tctgcgtgga gacccgggcc 360

gcctctgtgg gtttgcctag tgtttctctt gatctgccca ggctcagcat acaaaaagac 420

atacttacaa ttaaggctaa tacaactctt caaattactt gcaggggaca gagggacttg 480

gactggcttt ggcccaataa tcagagtggc agtgagcaaa gggtggaggt gactgagtgc 540

agcgatggcc tcttctgtaa gacactcaca attccaaaag tgatcggaaa tgacactgga 600

gcctacaagt gcttctaccg ggaaactgac ttggcctcgg tcatttatgt ctatgttcaa 660

gattacagat ctccatttat tgcttctgtt agtgaccaac atggagtcgt gtacattact 720

gagaacaaaa acaaaactgt ggtgattcca tgtctcgggt ccatttcaaa tctcaacgtg 780

tcactttgtg caagataccc agaaaagaga tttgttcctg atggtaacag aatttcctgg 840

gacagcaaga agggctttac tattcccagc tacatgatca gctatgctgg catggtcttc 900

tgtgaagcaa aaattaatga tgaaagttac cagtctatta tgtacatagt tgtcgttgta 960

gggtatagga tttatgatgt ggttctgagt ccgtctcatg gaattgaact atctgttgga 1020

gaaaagcttg tcttaaattg tacagcaaga actgaactaa atgtggggat tgacttcaac 1080
tgggaatacc cttcttcgaa gcatcagcat aagaaacttg taaaccgaga cctaaaaacc 1140
cagtctggga gtgagatgaa gaaatttttg agcaccttaa ctatagatgg tgtaacccgg 1200
agtgaccaag gattgtacac ctgtgcagca tccagtgggc tgatgaccaa gaagaacagc 1260
acatttgtca gggtccatga aaaacctttt gttgcttttg gaagtggcat ggaatctctg 1320
gtggaagcca cggtggggga gcgtgtcaga atccctgcga agtaccttgg ttacccaccc 1380
ccagaaataa aatggtataa aaatggaata ccccttgagt ccaatcacac aattaaagcg 1440
gggcatgtac tgacgattat ggaagtgagt gaaagagaca caggaaatta cactgtcatc 1500
cttaccaatc ccatttcaaa ggagaagcag agccatgtgg tctctctggt tgtgtatgtc 1560
ccaccccaga ttggtgagaa atctctaatc tctcctgtgg attcctacca gtacggcacc 1620
actcaaacgc tgacatgtac ggtctatgcc attcctcccc cgcatcacat ccactggtat 1680
tggcagttgg aggaagagtg cgccaacgag cccagccaag ctgtctcagt gacaaaccca 1740
taccttgtg aagaatggag aagtgtggag gacttccagg gaggaaataa aattgaagtt 1800
aataaaaatc aatttgctct aattgaagga aaaaacaaaa ctgtaagtac ccttgttatc 1860
caagcggcaa atgtgtcagc tttgtacaaa tgtgaagcgg tcaacaaagt cgggagagga 1920
gagagggtga tctccttcca cgtgaccagg ggtcctgaaa ttactttgca acctgacatg 1980
cagcccactg agcaggagag cgtgtctttg tggtgcactg cagacagatc tacgtttgag 2040
aacctcacat ggtacaagct tggcccacag cctctgccaa tccatgtggg agagttgccc 2100
acacctgttt gcaagaactt ggatactctt tggaaattga atgccaccat gttctctaat 2160
agcacaaatg acattttgat catggagctt aagaatgcat ccttgcagga ccaaggagac 2220
tatgtctgcc ttgctcaaga caggaagacc aagaaaagac attgcgtggt caggcagctc 2280
acagtcctag agcgtgtggc acccacgatc acaggaaacc tggagaatca gacgacaagt 2340
attggggaaa gcatcgaagt ctcatgcacg gcatctggga atccccctcc acagatcatg 2400
tggtttaaag ataatgagac ccttgtagaa gactcaggca ttgtattgaa ggatgggaac 2460
cggaacctca ctatccgcag agtgaggaag gaggacgaag gcctctacac ctgccaggca 2520
tgcagtgttc ttggctgtgc aaaagtggag gcatttttca taatagaagg tgcccaggaa 2580
aagacgaact tggaaatcat tattctagta ggcacggcgg tgattgccat gttcttctgg 2640
ctacttcttg tcatcatcct acggaccgtt aagcgggcca atggagggga actgaagaca 2700
ggctacttgt ccatcgtcat ggatccagat gaactcccat tggatgaaca ttgtgaacga 2760
ctgccttatg atgccagcaa atgggaattc cccagagacc ggctgaagct aggtaagcct 2820

cttggccgtg gtgcctttgg ccaagtgatt gaagcagatg cctttggaat tgacaagaca 2880
gcaacttgca ggacagtagc agtcaaaatg ttgaaagaag gagcaacaca cagtgagcat 2940
cgagctctca tgtctgaact caagatcctc attcatattg gtcaccatct caatgtggtc 3000
aaccttctag gtgcctgtac caagccagga gggccactca tggtgattgt ggaattctgc 3060
aaatttggaa acctgtccac ttacctgagg agcaagagaa atgaatttgt cccctacaag 3120
accaaagggg cacgattccg tcaagggaaa gactacgttg gagcaatccc tgtggatctg 3180
aaacggcgct tggacagcat caccagtagc cagagctcag ccagctctgg atttgtggag 3240
gagaagtccc tcagtgatgt agaagaagag gaagctcctg aagatctgta taaggacttc 3300
ctgaccttgg agcatctcat ctgttacagc ttccaagtgg ctaagggcat ggagttcttg 3360
gcatcgcgaa agtgtatcca cagggacctg gcggcacgaa atatcctctt atcggagaag 3420
aacgtggtta aaatctgtga ctttggcttg gcccgggata tttataaaga tccagattat 3480
gtcagaaaag gagatgctcg cctccctttg aaatggatgg ccccagaaac aatttttgac 3540
agagtgtaca caatccagag tgacgtctgg tcttttggtg ttttgctgtg ggaaatattt 3600
tccttaggtg cttctccata tcctggggta aagattgatg aagaattttg taggcgattg 3660
aaagaaggaa ctagaatgag ggcccctgat tatactacac cagaaatgta ccagaccatg 3720
ctggactgct ggcacgggga gcccagtcag agacccacgt tttcagagtt ggtggaacat 3780
ttgggaaatc tcttgcaagc taatgctcag caggatggca aagactacat tgttcttccg 3840
atatcagaga ctttgagcat ggaagaggat tctggactct ctctgcctac ctcacctgtt 3900
tcctgtatgg aggaggagga agtatgtgac cccaaattcc attatgacaa cacagcagga 3960
atcagtcagt atctgcagaa cagtaagcga aagagccggc ctgtgagtgt aaaaacattt 4020
gaagatatcc cgttagaaga accagaagta aaagtaatcc cagatgacaa ccagacggac 4080
agtggtatgg ttcttgcctc agaagagctg aaaactttgg aagacagaac caaattatct 4140
ccatcttttg gtggaatggt gcccagcaaa agcagggagt ctgtggcatc tgaaggctca 4200
aaccagacaa gcggctacca gtccggatat cactccgatg acacagacac caccgtgtac 4260
tccagtgagg aagcagaact tttaaagctg atagagattg gagtgcaaac cggtagcaca 4320
gcccagattc tccagcctga ctcggggacc acactgagct ctcctcctgt ttaaaaggaa 4380
gcatccacac cccaactccc ggacatcaca tgagaggtct gctcagattt tgaagtgttg 4440
ttctttccac cagcaggaag tagccgcatt tgattttcat ttcgacaaca gaaaaaggac 4500
ctcggactgc agggagccag tcttctaggc atatcctgga agaggcttgt gacccaagaa 4560
tgtgtctgtg tcttctccca gtgttgacct gatcctcttt tttcattcat ttaaaaagca 4620
ttatcatgcc cctgctgcgg gtctcaccat gggtttagaa caaagagctt caagcaatgg 4680

ccccatcctc aaagaagtag cagtacctgg ggagctgaca cttctgtaaa actagaagat 4740

aaaccaggca acgtaagtgt tcgaggtgtt gaagatggga aggatttgca gggctgagtc 4800

tatccaagag gctttgttta ggacgtgggt cccaagccaa gccttaagtg tggaattcgg 4860

attgatagaa aggaagacta acgttacctt gctttggaga gtactggagc ctgcaaatgc 4920

attgtgtttg ctctggtgga ggtgggcatg gggtctgttc tgaaatgtaa agggttcaga 4980

cggggtttct ggttttagaa ggttgcgtgt tcttcgagtt gggctaaagt agagttcgtt 5040

gtgctgtttc tgactcctaa tgagagttcc ttccagaccg ttagctgtct ccttgccaag 5100

ccccaggaag aaaatgatgc agctctggct ccttgtctcc caggctgatc ctttattcag 5160

aataccacaa agaaaggaca ttcagctcaa ggctccctgc cgtgttgaag agttctgact 5220

gcacaaacca gcttctggtt tcttctggaa tgaataccct catatctgtc ctgatgtgat 5280

atgtctgaga ctgaatgcgg gaggttcaat gtgaagctgt gtgtggtgtc aaagtttcag 5340

gaaggatttt accctttgt tcttcccccc gtccccaacc cactctcacc ccgcaaccca 5400

tcagtatttt agttatttgg cctctactcc agtaaacctg attgggtttg ttcactctct 5460

gaatgattat tagccagact tcaaaattat tttatagccc aaattataac atctattgta 5520

ttatttagac ttttaacata tagagctatt tctactgatt tttgcccttg ttctgtcctt 5580

tttttcaaaa aagaaaatgt gttttttgtt tggtaccata gtgtgaaatg ctgggaacaa 5640

tgactataag acatgctatg gcacatatat ttatagtctg tttatgtaga aacaaatgta 5700

atatattaaa gccttatata taatgaactt tgtactattc acattttgta tcagtattat 5760

gtagcataac aaaggtcata atgctttcag caattgatgt cattttatta aagaacattg 5820

aaaaacttga 5830

<210> 4
<211> 3334
<212> DNA
<213> Homo Sapiens

<400> 4
gcacgagccc cgggctgccg gcgcgggcgc cgcggcacgt ccacaggctg ggtcgcgagg 60

tggcgatcgc tgagaggcag gagggccgag gcgggcctgg gaggcggccc ggaggtgggg 120

cgccgctggg gccggcccgc acgggcttca tctgagggcg cacggcccgc gaccgagcgt 180

gcggactggc ctcccaagcg tggggcgaca agctgccgga gctgcaatgg gccgcggctg 240

gggattcttg tttggcctcc tgggcgccgt gtggctgctc agctcgggcc acggagagga 300

gcagcccccg gagacagcgg cacagaggtg cttctgccag gttagtggtt acttggatga 360

ttgtacctgt gatgttgaaa ccattgatag atttaataac tacaggcttt tcccaagact 420

acaaaaactt cttgaaagtg actactttag gtattacaag gtaaacctga agaggccgtg 480

tcctttctgg aatgacatca gccagtgtgg aagaagggac tgtgctgtca aaccatgtca 540

atctgatgaa gttcctgatg gaattaaatc tgcgagctac aagtattctg aagaagccaa 600

taatctcatt gaagaatgtg aacaagctga acgacttgga gcagtggatg aatctctgag 660

tgaggaaaca cagaaggctg ttcttcagtg gaccaagcat gatgattctt cagataactt 720

ctgtgaagct gatgacattc agtcccctga agctgaatat gtagatttgc ttcttaatcc 780

tgagcgctac actggttaca agggaccaga tgcttggaaa atatggaatg tcatctacga 840

agaaaactgt tttaagccac agacaattaa aagaccttta aatcctttgg cttctggtca 900

agggacaagt gaagagaaca ctttttacag ttggctagaa ggtctctgtg tagaaaaaag 960

agcattctac agacttatat ctggcctaca tgcaagcatt aatgtgcatt tgagtgcaag 1020

atatctttta caagagacct ggttagaaaa gaaatgggga cacaacatta cagaatttca 1080

acagcgattt gatggaattt tgactgaagg agaaggtcca agaaggctta agaacttgta 1140

ttttctctac ttaatagaac taagggcttt atccaaagtg ttaccattct tcgagcgccc 1200

agattttcaa ctctttactg gaaataaaat tcaggatgag gaaaacaaaa tgttacttct 1260

ggaaatactt catgaaatca agtcatttcc tttgcatttt gatgagaatt cattttttgc 1320

tggggataaa aaagaagcac acaaactaaa ggaggacttt cgactgcatt ttagaaatat 1380

ttcaagaatt atggattgtg ttggttgttt taaatgtcgt ctgtggggaa agcttcagac 1440

tcagggtttg ggcactgctc tgaagatctt attttctgag aaattgatag caaatatgcc 1500

agaaagtgga cctagttatg aattccatct aaccagacaa gaaatagtat cattattcaa 1560

cgcatttgga agaatttcta caagtgtgaa agaattagaa aacttcagga acttgttaca 1620

gaatattcat taaagaaaac aagctgatat gtgcctgttt ctggacaatg gaggcgaaag 1680

agtggaattt cattcaaagg cataatagca atgacagtct taagccaaac attttatata 1740

aagttgcttt tgtaaaggag aattatattg ttttaagtaa acacattttt aaaaattgtg 1800

ttaagtctat gtataatact actgtgagta aaagtaatac tttaataatg tggtacaaat 1860

tttaaagttt aatattgaat aaaaggagga ttatcaaatt catatatgat aaaagtgaat 1920

gttctaagtc tctcaaacta gcgttttatg taataatatg taatataaat aaaactatgg 1980

taaatgtgac aagcatttaa taggaaaatg ctaaggaggc ctcataaatg acccataatt 2040

accaacgtag aatttttcag tacatttagg gttgctggat ttagcaaata aaaataaaga 2100

ttgcccagtt agatttgaat ttcagataaa caattagttt tttaatattt tacatggaat 2160

atttggaaaa tacttatact aaaaaattat ttgtttgaaa ttcacattta actgggagtc 2220

ttgtatttta tctggcaatc ctaaaataca ttggtatgaa acaaatcact tttagaagta 2280

tattgctatt ttgattgggt tgtttttgtg tgtagaaacg tacaataaca actcaaaggc 2340

acaggagatt tctaaacatt gtgaaaagtt gaatagatta tatatttatt ctcataatac 2400

tttcactaat actaaataaa atttggggaa cacttttttat ttttatataa tttccaattt 2460

acagaaaagt ttcaaaaata gtacaaagag ctctcttacc cagattcact aattgttcat 2520

acgtgcttta tctttcatgc tttctctgta cacacacaca cacacacaaa tttttcctca 2580

atcatttgaa agtcagttat aggcatcatg ccccttaaac cctaaatact tcagtgtgta 2640

atactgaata attactaaaa atgattttct cagaaaaaaa aactcccaca attctggaac 2700

tataatactg taagccttag aataaataat actttcaagt tccaatctaa agttcttttt 2760

gagttttgtt gcccgtttta tgcttgatgt gtatagtaat agggtaggct atttatttta 2820

ttaaaatttt ttttagagac aaggttttgc tgtgttgccc aagctggaac ttgaacgact 2880

gggctgaagt gatcttccca cctcagcctc ccaagtagct gggaatacag gtgtctgcca 2940

ccatacccag tttcattttt gtttttttata cccgaagttc atttcctttg tctccctaaa 3000

actgaactgt aattttggga ggttttcatt agtggaagct cttcatttat aaagctattt 3060

gaaggggttt aggaatttat atcacatggt aattgtagag aaaaagaagc tatatacctc 3120

aaaatcgtgc cctctttaca tatgtcttat caggtataac atgttgaaat gtcacattag 3180

tagtaaagtg gggtttattt atatagtggt taagaaatgt cagtttacac tgctgtatac 3240

ttcttcttct gtgtccctaa ggcctggtac agtgccaagc acatacttgg tatccaataa 3300

atatttgttg gatgaaaaaa aaaaaaaaaa aaaa 3334

<210> 5
<211> 840
<212> DNA
<213> Homo Sapiens

<400> 5
ccacgtccgg ggtgccgagc caactttcct gcgtccatgc agccccgccg gcaacggctg 60

cccgctccct ggtccgggcc caggggcccg cgccccaccg ccccgctgct cgcgctgctg 120

ctgttgctcg ccccggtggc ggcgcccgcg gggtccgggg gccccgacga ccctgggcag 180

cctcaggatg ctggggtccc gcgcaggctc ctgcagcaga aggcgcgcgc ggcgcttcac 240

ttcttcaact tccggtccgg ctcgcccagc gcgctgcgag tgctggccga ggtgcaggag 300

ggccgcgcgt ggattaatcc aaaagaggga tgtaaagttc acgtggtctt cagcacagag 360

cgctacaacc cagagtcttt acttcaggaa ggtgagggac gtttggggaa atgttctgct 420

cgagtgtttt tcaagaatca gaaacccaga ccaaccatca atgtaacttg tacacggctc 480

atcgagaaaa agaaaagaca acaagaggat tacctgcttt acaagcaaat gaagcaactg 540

aaaaacccct tggaaatagt cagcatacct gataatcatg gacatattga tccctctctg 600

agactcatct gggatttggc tttccttgga agctcttacg tgatgtggga aatgacaaca 660

caggtgtcac actactactt ggcacagctc actagtgtga ggcagtgggt aagaaaaacc 720

tgaaaattaa cttgtgccac aagagttaca atcaaagtgg tctccttaga ctgaattcat 780

gtgaacttct aatttcatat caagagttgt aatcacattt atttcaataa atatgtgagt 840

<210> 6
<211> 3314
<212> DNA
<213> Homo Sapiens

<400> 6
ggaggcaggc ggtgccgcgg cgccgggacc cgactcatcc ggtgcttgcg tgtggtggtg 60

agcgcagcgc cgaggatgag gaggtgcaac agcggctccg ggccgccgcc gtcgctgctg 120

ctgctgctgc tgtggctgct cgcggttccc ggcgctaacg cggccccgcg gtcggcgctc 180

tattcgcctt ccgacccgct gacgctgctg caggcggaca cggtgcgcgg cgcggtgctg 240

ggctcccgca gcgcctgggc cgtggagttc ttcgcctcct ggtgcggcca ctgcatcgcc 300

ttcgccccga cgtggaaggc gctggccgaa gacgtcaaag cctggaggcc ggccctgtat 360

ctcgccgccc tggactgtgc tgaggagacc aacagtgcag tctgcagaga cttcaacatc 420

cctggcttcc cgactgtgag gttcttcaag gcctttacca agaacggctc gggagcagta 480

tttccagtgg ctggtgctga cgtgcagacg ctgcgggaga ggctcattga cgccctggag 540

tcccatcatg acacgtggcc cccagcctgt cccccactgg agcctgccaa gctggaggag 600

attgatggat tctttgcgag aaataacgaa gagtacctgg ctctgatctt tgaaaaggga 660

ggctcctacc tgggtagaga ggtggctctg gacctgtccc agcacaaagg cgtggcggtg 720

cgcagggtgc tgaacacaga ggccaatgtg gtgagaaagt ttggtgtcac cgacttcccc 780

tcttgctacc tgctgttccg gaatggctct gtctcccgag tccccgtgct catggaatcc 840

aggtccttct ataccgctta cctgcagaga ctctctgggc tcaccaggga ggctgcccag 900

accacagttg caccaaccac tgctaacaag atagctccca ctgtttggaa attggcagat 960

cgctccaaga tctacatggc tgacctggaa tctgcactgc actacatcct gcggatagaa 1020

gtgggcaggt tcccggtcct ggaagggcag cgcctggtgg ccctgaaaaa gtttgtggca 1080

gtgctggcca agtatttccc tggccggccc ttagtccaga acttcctgca ctccgtgaat 1140

gaatggctca agaggcagaa gagaaataaa attccctaca gtttctttaa aactgccctg 1200

gacgacagga aagagggtgc cgttcttgcc aagaaggtga actggattgg ctgccagggg 1260

agtgagccgc atttccgggg ctttccctgc tccctgtggg tcctcttcca cttcttgact 1320
gtgcaggcag ctcggcaaaa tgtagaccac tcacaggaag cagccaaggc caaggaggtc 1380
ctcccagcca tccgaggcta cgtgcactac ttcttcggct gccgagactg cgctagccac 1440
ttcgagcaga tggctgctgc ctccatgcac cgggtgggga gtcccaacgc cgctgtcctc 1500
tggctctggt ctagccacaa cagggtcaat gctcgccttg caggtgcccc cagcgaggac 1560
ccccagttcc ccaaggtgca gtggccaccc cgtgaacttt gttctgcctg ccacaatgaa 1620
cgcctggatg tgcccgtgtg ggacgtggaa gccaccctca acttcctcaa ggcccacttc 1680
tccccaagca acatcatcct ggacttccct gcagctgggt cagctgcccg gagggatgtg 1740
cagaatgtgg cagccgcccc agagctggcg atgggagccc tggagctgga aagccggaat 1800
tcaactctgg accctgggaa gcctgagatg atgaagtccc ccacaaacac caccccacat 1860
gtgccggctg agggacctga ggcaagtcga cccccgaagc tgcaccctgg cctcagagct 1920
gcaccaggcc aggagcctcc tgagcacatg gcagagcttc agaggaatga gcaggagcag 1980
ccgcttgggc agtggcactt gagcaagcga gacacagggg ctgcattgct ggctgagtcc 2040
agggctgaga agaaccgcct ctggggccct ttggaggtca ggcgcgtggg ccgcagctcc 2100
aagcagctgg tcgacatccc tgagggccag ctggaggccc gagctggacg gggccgaggc 2160
cagtggctgc aggtgctggg agggggcttc tcttacctgg acatcagcct ctgtgtgggg 2220
ctctattccc tgtccttcat gggcctgctg gccatgtaca cctacttcca ggccaagata 2280
agggccctga agggccatgc tggccaccct gcagcctgaa ccacctgggg aggaggcggg 2340
agagggagct gccatctcta ggcacctcaa gcccccctgac cccattcccc cccctcccac 2400
cccttgctcc ttgtctggcc tagaagtgtg ggaaattcag gaaaacgagt tgctccagtg 2460
aagcttcttg gggttgctag gacagagagc tcctttgaca caaaagacag gagcagggtc 2520
caggttcccc tgctgtgcag ggagggcagc cccgggcagt gggcataggg cagctcagtc 2580
cctggcctct tagcaccaca ttcctgtttt tcagcttatt tgaagtcctg cctcattctc 2640
actggagcct cagtctctcc tgcttggtct tggccctcaa ctggggcaag tgaagccaga 2700
ggagggtccc ccagctgggt gggctggaat ggaactcctc actagctgct ggggctccgc 2760
ccaccctgct cccttccgga caatgaagaa gcctttgcac cctgggagga aggaccaccc 2820
cgggccctct atgcctggcc agcctccagc tcctcagacc tcctgggtgg ggtttggctt 2880
cagggtgggg tttggaagct tctggaagtc gtgctggtct cccaggtgag gcaagccatg 2940
gttgctgggc tgtagggtga gtggcttgct tggtgggacc tgacgagttg gtggcatggg 3000
aaggatgtgg gtctctagtg ccttgccctg gcttagctgc aggagaagat ggctgctttc 3060
acttcccccc attgagctct gctccctctg agcctggtct tttgtccttt tttattttgg 3120

tctccaagat gaatgctcat ctttggaggg tgccaggtag aagctaggga ggggagtgtc 3180

ttctctctcc aggtttcacc ttccagtgtg cagaagttag aagggtctgg cgggggcagt 3240

gccttacaca tgcttgattc ccacgctacc ccctgccttg ggaggtgtgt ggaataaatt 3300

atttttgtta aggc 3314

<210> 7
<211> 4020
<212> DNA
<213> Homo Sapiens

<400> 7
ggcacgaggg tggagccgag cggtgcggag cagatctggt ggttctccgg agagcagctt 60

ccttgggtgt tacatgagcc aagccctcac tgtacagaag agtgagagct gaaacctgtt 120

ccctgagctg atcagaagga catcccttgg cccctccatc tgggctcctg tggataggag 180

gggctgggtg agcaggccag ctgggctatg gtgtggtgcc tcggcctggc cgtcctcagc 240

ctggtcatca gccagggggc tgacggtcga gggaagcctg aggtggtatc ggtggtgggc 300

cgggctgagg agagtgtggt gctgggctgt gacctgctgc ccccggccgg ccggcccccc 360

ctgcatgtca tcgagtggct gcgctttgga ttcctgcttc ccatcttcat ccagttcggc 420

ctctactctc cccgaattga ccctgattac gtgggacgag tccggctgca gaaggggccc 480

tctctccaga ttgagggtct ccgggtggaa gaccagggct ggtacgagtg ccgcgtgttc 540

ttcctggacc agcacatccc tgaagacgat tttgctaacg gctcctgggt gcatctgaca 600

gtcaattcac cccctcaatt ccaggagaca cctcctgctg tgttggaagt gcaggaactg 660

gagcctgtga ccctgcgttg tgtggcccgt ggcagccccc tgcctcatgt gacgtggaag 720

ctccgaggaa aggaccttgg ccagggccag ggccaggtgc aagtgcagaa cgggacgctg 780

cggatccgcc gggtagagcg aggcagctct ggggtctaca cctgccaagc ctccagcact 840

gagggcagcg ccacccacgc cacccagctg ctagtgctag gacccccagt catcgtggtg 900

ccccccaaga acagcacagt caatgcctcc caggatgttt cattggcctg ccatgctgag 960

gcataccctg ctaacctcac ctacagctgg ttccaggaca acatcaatgt cttccacatt 1020

agccgcctgc agccccgggt gcagatcctg gtggacggga gcctgcggct gctggccacc 1080

cagcctgatg atgccggctg ctacacctgt gtgcccagca atggcctcct gcatccaccc 1140

tcagcctctg cctacctcac tgtgctctgc atgccggggg tgatccgctg cccggttcgt 1200

gccaaccccc cactgctctt tgtcagctgg accaaggatg gaaaggccct gcagctggac 1260

aagttccctg gctggtccca gggcacagaa ggctcactga tcatcgccct ggggaacgag 1320

gatgccctgg gagaatactc ctgcaccccc tacaacagtc ttggtaccgc cgggccctct 1380

cctgtgaccc gcgtgctgct caaggctccc ccagctttta tagagcggcc caaggaagaa 1440

tatttccaag aagtagggcg ggagctgctc atcccctgct ccgcccaagg ggaccctcct 1500

cctgttgtct cttggaccaa ggtgggccgg gggctgcaag gccaggccca ggtggacagc 1560

aacagcagcc tcatcctgcg accattgacc aaggaggccc acgggcactg ggaatgcagt 1620

gccagcaatg ctgtggcccg agtggccacc tccacgaacg tctacgtgct gggcactagc 1680

cctcatgttg tcaccaatgt gtccgtggtg gctttgccca agggtgccaa tgtctcctgg 1740

gagcctggct ttgatggtgg ttatctgcag agattcagtg tctggtacac cccactggcc 1800

aagcgtcctg accgaatgca ccatgactgg gtgtccttgg cagtgcctgt gggggctgct 1860

cacctcctag tgccagggct gcagccccac acccagtacc agttcagcgt gctagctcag 1920

aacaagctgg ggagtggtcc cttcagcgaa atcgtcttgt ctgctccgga agggcttcct 1980

accacgccag ctgcacccgg gcttcccccа acagagatac cgcctcccct gtcccctccg 2040

cggggtctgg tggcagtgag gacaccccgg ggggtactcc tgcattggga tcccccagag 2100

ctggtcccta agagactgga tggctacgtc ttggaaggcc ggcaaggctc ccagggctgg 2160

gaggtgctgg acccggctgt ggcaggcaca gaaacagagc tgctggtgcc aggcctcatc 2220

aaggatgttc tctacgagtt ccgcctcgtg gccttcgcgg gcagcttcgt cagcgacccc 2280

agcaacacgg ccaacgtctc cacttccggt ctggaggtct acccttcgcg cacgcagctg 2340

ccgggcctcc tgcctcagcc cgtgctggcc ggcgtggtgg gcggagtctg ctttctggga 2400

gtggccgtcc ttgtgagcat cctggccggc tgcctcctga accggcgcag ggctgcccgc 2460

cgccgccgca agcgcctccg ccaagatcca cctcttatct tctctccgac cgggaagtca 2520

gctgcaccct ctgctctggg ctcaggcagt cctgacagcg tggcgaagct gaagctccag 2580

ggatccccag tccccagcct gcgccagagt ctgctctggg gggatcctgc cggaactccc 2640

agcccccacc cggatcctcc atctagccgg ggacccttac ctctggagcc catttgccgg 2700

ggcccagacg ggcgctttgt gatggggccc actgtggcgg ccccccagga aaggtcaggc 2760

cgggagcagg cagaacctcg gactccagcc cagcgtctgg cccggtcctt tgactgtagc 2820

agcagcagcc ccagtggggc accccagccc ctctgcattg aagacatcag ccctgtggca 2880

ccccctccag cagccccacc cagtcccttg ccaggtcctg gacccctgct ccagtacctg 2940

agcctgccct tcttccgaga gatgaatgtg gatggggact ggcccccgct tgaggagccc 3000

agccctgctg cacccccaga ttacatggat acccggcgct gtcccacctc atctttcctt 3060

cgttctccag aaacccctcc tgtatccccc agggaatcac ttcctggggc tgtggtaggg 3120

gctggggcca ctgcagagcc cccttacaca gccctggctg actggacact gagggagcgg 3180

ctgctgccag gccttctccc tgctgcccct cgaggcagcc tcaccagcca gagcagcggg 3240

cgaggcagcg cttcgttcct gcggcccccc tccacagccc cctctgcagg aggcagctac 3300

ctcagccctg ctccaggaga caccagcagc tgggccagtg gccctgagag atggccccga 3360

agggagcatg tggtgacagt cagcaagagg aggaacacat ctgtggacga gaactatgag 3420

tgggactcag aattccctgg ggacatggaa ttgctggaga ctttgcacct gggcttggcc 3480

agctcccggc tcagacctga agctgagaca gagctaggtg tgaagactcc agaggagggc 3540

tgcctcctga acactgccca tgttactggc cctgaggccc gctgtgctgc ccttcgggag 3600

gaattcctgg ccttccgccg ccgccgagat gctactaggg ctcggctacc agcctatcga 3660

cagccagtcc cccaccccga acaggccact ctgctgtgaa catccctaat gtgaggctgt 3720

gaaaaggcat atggacctgc aaaggaggcc cccaaccaga cagacctagt ttcaaacgag 3780

ggcactgccc ctgcctgccc ctttggtgcc caggcacaga ccctgatagt gggtttgggt 3840

caccttggta tggaatgtat gtgctgaccc cctaggtgag tctggggatt ggaacaggga 3900

tcttaggtct gcctctctct ctctctctct ctctctctct ctctctctgt gtgtgtgtgt 3960

gtgtgtgaag ttttttacag gtgaataaac aaagtttgaa agaaaaaaaa aaaaaaaaaa 4020

<210> 8
<211> 1284
<212> DNA
<213> Homo Sapiens

<400> 8
ggcacgaggg tctccgcctg caggtgcaga catctggagg agagagtcgg agagcagaaa 60

ccacttggct cccagacaat tcccctacag gctttgggcc tggaattgag gagaaagtga 120

gctaagttgg ggtggggtga gtccaaagaa gcacgggctg ggccaagcta agctgctctg 180

ggctgggctg atccctcccc actcaggggc gggaccccag gaggagggag aggacagagc 240

cactgcagag gaccagactg ggaaaacaac gatatggcag gagccagtct tggggcccgc 300

ttctaccggc agatcaaaag acatccgggg atcatcccga tgatcggctt aatctgcctg 360

ggcatgggca gcgctgcgct ttacttgctg cgactcgccc ttcgcagccc cgacgtctgc 420

tgggacagaa agaacaaccc ggagccctgg aaccgcctga gccccaatga ccaatacaag 480

ttccttgcag tttccactga ctataagaag ctgaagaagg accggccaga cttctaagcc 540

aggctgggct gccagtgcca tgcaagccac agccagccag cccatccact tcttccactc 600

ctccccgcag gccccaaggc atcactccgg ccaccctgtc ccgctactgc ttacacaggc 660

cgggttccca cgcagagggg aggctgctcc acccctactc tcctcccttg ctcccagcag 720

cggaagcgcc tctgaccctt ggcttgagtc ccacgtgggg gaggaggagg caggcagcac 780

cagcaggggt ccaccaagag cccagaccag cccctctgcc ctcctacccg ggcctcgaag 840

ggtgtggcac aggctacgtg ttgagcgtgg cctacgtgag ccaacaagaa gcaggggcct 900

ctgagtgcca agcgacgtgg cgggctccac gttagcccag gctctgagag ccagcccagg 960

ggcggcgctg ctcagcttgg gctggtccag ggcctgccca ggctggggca cctttgcctc 1020

ctgaggcgca gcgcactcct cccctgccca agcctactgc ctcccgctgc cgccagtacc 1080

ccctccagcc ccacacctgg gcctcccccc gccactcccc tcccttgctc ccctctgtcc 1140

ccagggatca aacagaagca gccgtgggca aaatacaatt tcatttaaca aattgaaaaa 1200

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1260

aaaaaaaaaa aaaaaaaaaa aaaa 1284

<210> 9
<211> 4165
<212> DNA
<213> Homo Sapiens

<220>
<221> misc_feature
<222> (4076)..(4076)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (4091)..(4091)
<223> n is a, c, g, or t

<400> 9
cgggctactt tgaaaggaca accatttttc tttccgctaa tttataatgg ttttgaagtg 60

gttgttcatt ctcaaacata gacttttaaa tgttaggtct ttcctataac tctttgttat 120

tggaagtttc aaggatttgg acactcaatt aaggattctg tcctctcctc attcctttgg 180

ttttggccca aatgattatg tttcctcttt ttgggaagat ttctctgggt attttgatat 240

ttgtcctgat agaaggagac tttccatcat taacagcaca aacctactta tctatagagg 300

agatccaaga acccaagagt gcagtttctt ttctcctgcc tgaagaatca acagaccttt 360

ctctagctac caaaaagaaa cagcctctgg accgcagaga aactgaaaga cagtggttaa 420

tcagaaggcg gagatctatt ctgtttccta atggagtgaa aatctgccca gatgaaagtg 480

ttgcagaggc tgtggcaaat catgtgaagt attttaaagt ccgagtgtgt caggaagctg 540

tctgggaagc cttcaggact ttttgggatc gacttcctgg gcgtgaggaa tatcattact 600

ggatgaattt gtgtgaggat ggagtcacaa gtatatttga aatgggcaca aattttagtg 660

aatctgtgga acatagaagc ttaatcatga agaaactgac ttatgcaaag gaaactgtaa 720

gcagctctga actgtcttct ccagttcctg ttggtgatac ttcaacattg ggagacacta 780

ctctcagtgt tccacatcca gaggtggacg cctatgaagg tgcctcagag agcagcttgg 840

aaaggccaga ggagagtatt agcaatgaaa ttgagaatgt gatagaagaa gccacaaaac 900

cagcaggtga acagattgca gaattcagta tccacctttt ggggaagcag tacagggaag 960

aactacagga ttcctccagc tttcaccacc agcaccttga agaagaattt atttcagagg 1020

ttgaaaatgc atttactggg ttaccaggct acaaggaaat tcgtgtactt gaatttaggt 1080

cccccaagga aaatgacagt ggcgtagatg tttactatgc agttaccttc aatggtgagg 1140

ccatcagcaa taccacctgg gacctcatta gccttcactc caacaaggtg gaaaaccatg 1200

gccttgtgga actggatgat aaacccactg ttgtttatac aatcagtaac ttcagagatt 1260

atattgctga gacattgcag cagaattttt tgctggggaa ctcttccttg aatccagatc 1320

ctgattccct gcagcttatc aatgtgagag gagttttgcg tcaccaaact gaagatctag 1380

tttggaacac ccaaagttca agtcttcagg caacgccgtc atctattctg gataatacct 1440

ttcaagctgc atggccctca gcagatgaat ccatcaccag cagtattcca ccacttgatt 1500

tcagctctgg tcctccctca gccactggca gggaactctg gtcagaaagt cctttgggtg 1560

atttagtgtc tacacacaaa ttagcctttc cctcgaagat gggcctcagc tcttccccag 1620

aggttttaga ggttagcagc ttgactcttc attctgtcac cccggcagtg cttcagactg 1680

gcttgcctgt ggcttctgag gaaaggactt ctggatctca cttggtagaa gatggattag 1740

ccaatgttga agagtcagaa gattttcttt ctattgattc attgccttca agttcattca 1800

ctcaacctgt gccaaaagaa acaataccat ccatggaaga ctctgatgtg tccttaacat 1860

cttcaccata tctgacctct tctatacctt ttggcttgga ctccttgacc tccaaagtca 1920

aagaccaatt aaaagtgagc cctttcctgc cagatgcatc catggaaaaa gagttaatat 1980

ttgacggtgg tttaggttca gggtctgggc aaaaggtaga tctgattact tggccatgga 2040

gtgagacttc atcagagaag agcgccgaac cactgtccaa gccgtggctt gaagatgatg 2100

attcactttt gccagctgag attgaagaca agaaactagt tttagttgac aaaatggatt 2160

ccacagacca aattagtaag cactcaaaat atgaacatga tgacagatcc acacactttc 2220

cagaggaaga gcctcttagt gggcctgctg tgcccatctt cgcagatact gcagctgaat 2280

ctgcgtctct aaccctcccc aagcacatat cagaagtacc tggtgttgat gattgctcag 2340

ttaccaaagc acctcttata ctgacatctg tagcaatctc tgcctctact gataaatcag 2400

atcaggcaga tgccatccta agggaggata tggaacaaat tactgagtca tccaactatg 2460

aatggtttga cagtgaggtt tcaatggtaa agccagatat gcaaactttg tggactatat 2520

tgccagaatc agagagagtt tggacaagaa cttcttccct agagaaattg tccagagaca 2580

tattggcaag tacaccacag agtgctgaca ggctctggtt atctgtgaca cagtctacca 2640

aattgcctcc aaccacaatc tccaccctgc tagaggatga agtaattatg ggtgtacagg 2700

atatttcgtt agaactggac cggataggca cagattacta tcagcctgag caagtccaag 2760

agcaaaatgg caaggttggt agttatgtgg aaatgtcaac aagtgttcac tccacagaga 2820

tggttagtgt ggcttggccc acagaaggag gagatgactt gagttatacc cagacttcag 2880

gagctttggt ggttttcttc agcctccgag tgactaacat gatgttttca gaagatctgt 2940

ttaataaaaa ctccttggag tataaagccc tggagcaaag attcttagaa ttgctggttc 3000

cctatctcca gtcaaatctc acggggttcc agaacttaga aatcctcaac ttcagaaatg 3060

gcagcattgt ggtgaacagt cgaatgaagt ttgccaattc tgtccctcct aacgtcaaca 3120

atgcggtgta catgattctg gaagactttt gtaccactgc ctacaatacc atgaacttgg 3180

ctattgataa atactctctt gatgtggaat caggtgatga agccaaccct tgcaagtttc 3240

aggcctgtaa tgaattttca gagtgtctgg tcaacccctg gagtggagaa gcaaagtgca 3300

gatgcttccc tggatacctg agtgtggaag aacggccctg tcagagtctc tgtgacctac 3360

agcctgactt ctgcttgaat gatggaaagt gtgacattat gcctgggcac ggggccattt 3420

gtaggtgccg ggtgggtgag aactggtggt accgaggcaa gcactgtgag gaatttgtgt 3480

ctgagcccgt gatcataggc atcactattg cctccgtggt tggacttctt gtcatctttt 3540

ctgctatcat ctacttcttc atcaggactc ttcaagcaca ccatgacagg agtgaaagag 3600

agagtccctt cagtggctcc agcaggcagc ctgacagcct ctcatctatt gagaatgctg 3660

tgaagtacaa ccccgtgtat gaaagtcaca gggctggatg tgagaagtat gagggaccct 3720

atcctcagca tcccttctac agctctgcta gcggagacgt gattggtggg ctgagcagag 3780

aagaaatcag acagatgtat gagagcagtg agctttccag agaggaaatt caagagagaa 3840

tgagagtttt ggaactgtat gccaatgatc ctgagtttgc agcttttgtg agagagcaac 3900

aagtggaaga ggtttaacca aaactcctgt tctgaaactg attagaagcc tggagaagat 3960

ggagattact tgttacttat gtcatataat taacctggat tttaaacact gttggaagaa 4020

gagttttcta tgaaaaaatt aaatataggg cacactgttt ttttttcagc ttaagntttc 4080

agaatgtagt nagagatgtw mcatttttat ttctataaag actgaatgct gtgtttaaat 4140

aattgaaaac tacgttaaaa aaaaa 4165

<210> 10
<211> 1237
<212> DNA
<213> Homo Sapiens

<400> 10

gaattcggca cgaggcctcg tgccggggag caacctcagc ttctagtatc cagactccag 60

cgccgccccg ggcgcggacc ccaaccccga cccagagctt ctccagcggc ggcgcagcga 120

gcagggctcc ccgccttaac ttcctccgcg gggcccagcc accttcggga gtccgggttg 180

cccacctgca aactctccgc cttctgcacc tgccaccccc gagccagcgc gggcgcccga 240

gcgagtcatg gccaacgcgg ggctgcagct gttgggcttc attctcgcct tcctgggatg 300

gatcggcgcc atcgtcagca ctgccctgcc ccagtggagg atttactcct atgccggcga 360

caacatcgtg accgcccagg ccatgtacga ggggctgtgg atgtcctgcg tgtcgcagag 420

caccgggcag atccagtgca aagtctttga ctccttgctg aatctgagca gcacattgca 480

agcaacccgt gccttgatgg tggttggcat cctcctggga gtgatagcaa tctttgtggc 540

caccgttggc atgaagtgta tgaagtgctt ggaagacgat gaggtgcaga agatgaggat 600

ggctgtcatt gggggcgcga tatttcttct tgcaggtctg gctattttag ttgccacagc 660

atggtatggc aatagaatcg ttcaagaatt ctatgaccct atgaccccag tcaatgccag 720

gtacgaattt ggtcaggctc tcttcactgg ctgggctgct gcttctctct gccttctggg 780

aggtgccctа ctttgctgtt cctgtccccg aaaaacaacc tcttacccaa caccaaggcc 840

ctatccaaaa cctgcacctt ccagcgggaa agactacgtg tgacacagag gcaaaaggag 900

aaaatcatgt tgaaacaaac cgaaaatgga cattgagata ctatcattaa cattaggacc 960

ttagaatttt gggtattgta atctaaagta tgttattaca aaacaaacaa acaaacaaaa 1020

aacccatgtg ttaaaatact cagtgctaaa catggcttaa tcttatttta tcttctttcc 1080

tcaatatagg agggaagatt tttccatttg tattactgct tcccattgag taatcatact 1140

caaatggggg aaggggtgct ccttaaatat atatagatat gtatatatac atgtttttct 1200

attaaaaata gccagtaaaa aaaaaaaaaa aaaaaaa 1237

<210> 11
<211> 2010
<212> DNA
<213> Homo Sapiens

<400> 11
acagttgttg caaagtgctc agcactaagg gagccagcgc acagcacagc caggaaggcg 60

agcgagccca gccagcccag ccagcccagc cagcccggag gtcatttgat tgcccgcctc 120

agaacgatgg atctgcatct cttcgactac tcagagccag ggaacttctc ggacatcagc 180

tggccatgca acagcagcga ctgcatcgtg gtggacacgg tgatgtgtcc caacatgccc 240

aacaaaagcg tcctgctcta cacgctctcc ttcatttaca ttttcatctt cgtcatcggc 300

atgattgcca actccgtggt ggtctgggtg aatatccagg ccaagaccac aggctatgac 360

acgcactgct acatcttgaa cctggccatt gccgacctgt gggttgtcct caccatccca 420
gtctgggtgg tcagtctcgt gcagcacaac cagtggccca tgggcgagct cacgtgcaaa 480
gtcacacacc tcatcttctc catcaacctc ttcggcagca ttttcttcct cacgtgcatg 540
agcgtggacc gctacctctc catcacctac ttcaccaaca cccccagcag caggaagaag 600
atggtacgcc gtgtcgtctg catcctggtg tggctgctgg ccttctgcgt gtctctgcct 660
gacacctact acctgaagac cgtcacgtct gcgtccaaca atgagaccta ctgccggtcc 720
ttctaccccg agcacagcat caaggagtgg ctgatcggca tggagctggt ctccgttgtc 780
ttgggctttg ccgttccctt ctccattatc gctgtcttct acttcctgct ggccagagcc 840
atctcggcgt ccagtgacca ggagaagcac agcagccgga agatcatctt ctcctacgtg 900
gtggtcttcc ttgtctgctg gctgccctac cacgtggcgg tgctgctgga catcttctcc 960
atcctgcact acatcccttt cacctgccgg ctggagcacg ccctcttcac ggccctgcat 1020
gtcacacagt gcctgtcgct ggtgcactgc tgcgtcaacc ctgtcctcta cagcttcatc 1080
aatcgcaact acaggtacga gctgatgaag gccttcatct tcaagtactc ggccaaaaca 1140
gggctcacca agctcatcga tgcctccaga gtctcagaga cggagtactc tgccttggag 1200
cagagcacca aatgatctgc cctggagagg ctctgggacg ggtttacttg tttttgaaca 1260
gggtgatggg ccctatggtt ttctagagca aagcaaagta gcttcgggtc ttgatgcttg 1320
agtagagtga agaggggagc acgtgccccc tgcatccatt ctctctttct cttgatgacg 1380
cagctgtcat ttggctgtgc gtgctgacag ttttgcaaca ggcagagctg tgtcgcacag 1440
cagtgctgtg cgtcagagcc agctgaggac aggcttgcct ggacttctgt aagataggat 1500
tttctgtgtt tcctgaattt tttatatggt gatttgtatt taaattttaa gactttattt 1560
tctcactatt ggtgtacctt ataaatgtat ttgaaagtta aatatatttt aaatattgtt 1620
tgggaggcat agtgctgaca tatattcaga gtgttgtagt tttaaggtta gcgtgacttc 1680
agttttgact aaggatgaca ctaattgtta gctgttttga aattatatat atataaatat 1740
atataaatat ataaatatat gccagtcttg gctgaaatgt tttatttacc atagttttat 1800
atctgtgtgg tgttttgtac cggcacggga tatggaacga aaactgcttt gtaatgcagt 1860
ttgtgacatt aatagtattg taaagttaca ttttaaaata aacaaaaaac tgttctggac 1920
tgcaaatctg cacacacaac gaacagttgc atttcagaga gttctctcaa tttgtaagtt 1980
attttttttt aataaagatt tttgtttcct 2010

<210> 12
<211> 7318
<212> DNA
<213> Homo Sapiens

<400> 12
ctggctctta acggcgttta tgtcctttgc tgtctgaggg gcctcagctc tgaccaatct 60
ggtcttcgtg tggtcattag catgggcttc gtgagacaga tacagctttt gctctggaag 120
aactggaccc tgcggaaaag gcaaaagatt cgctttgtgg tggaactcgt gtggccttta 180
tctttatttc tggtcttgat ctggttaagg aatgccaacc cgctctacag ccatcatgaa 240
tgccatttcc ccaacaaggc gatgccctca gcaggaatgc tgccgtggct ccaggggatc 300
ttctgcaatg tgaacaatcc ctgttttcaa agccccaccc caggagaatc tcctggaatt 360
gtgtcaaact ataacaactc catcttggca agggtatatc gagattttca agaactcctc 420
atgaatgcac cagagagcca gcaccttggc cgtatttgga cagagctaca catcttgtcc 480
caattcatgg acaccctccg gactcacccg gagagaattg caggaagagg aatacgaata 540
agggatatct tgaaagatga agaaacactg acactatttc tcattaaaaa catcggcctg 600
tctgactcag tggtctacct tctgatcaac tctcaagtcc gtccagagca gttcgctcat 660
ggagtcccgg acctggcgct gaaggacatc gcctgcagcg aggccctcct ggagcgcttc 720
atcatcttca gccagagacg cggggcaaag acggtgcgct atgccctgtg ctccctctcc 780
cagggcaccc tacagtggat agaagacact ctgtatgcca acgtggactt cttcaagctc 840
ttccgtgtgc ttcccacact cctagacagc cgttctcaag gtatcaatct gagatcttgg 900
ggaggaatat tatctgatat gtcaccaaga attcaagagt ttatccatcg gccgagtatg 960
caggacttgc tgtgggtgac caggcccctc atgcagaatg gtggtccaga gacctttaca 1020
aagctgatgg gcatcctgtc tgacctcctg tgtggctacc ccgagggagg tggctctcgg 1080
gtgctctcct tcaactggta tgaagacaat aactataagg cctttctggg gattgactcc 1140
acaaggaagg atcctatcta ttcttatgac agaagaacaa catccttttg taatgcattg 1200
atccagagcc tggagtcaaa tcctttaacc aaaatcgctt ggagggcggc aaagcctttg 1260
ctgatgggaa aaatcctgta cactcctgat tcacctgcag cacgaaggat actgaagaat 1320
gccaactcaa cttttgaaga actggaacac gttaggaagt tggtcaaagc ctgggaagaa 1380
gtagggcccc agatctggta cttctttgac aacagcacac agatgaacat gatcagagat 1440
accctgggga acccaacagt aaaagacttt ttgaataggc agcttggtga agaaggtatt 1500
actgctgaag ccatcctaaa cttcctctac aagggccctc gggaaagcca ggctgacgac 1560
atggccaact tcgactggag ggacatattt aacatcactg atcgcaccct ccgcctggtc 1620
aatcaatacc tggagtgctt ggtcctggat aagtttgaaa gctacaatga tgaaactcag 1680
ctcacccaac gtgccctctc tctactggag gaaaacatgt tctgggccgg agtggtattc 1740
cctgacatgt atccctggac cagctctcta ccaccccacg tgaagtataa gatccgaatg 1800

gacatagacg tggtggagaa aaccaataag attaaagaca ggtattggga ttctggtccc 1860

agagctgatc ccgtggaaga tttccggtac atctggggcg ggtttgccta tctgcaggac 1920

atggttgaac aggggatcac aaggagccag gtgcaggcgg aggctccagt tggaatctac 1980

ctccagcaga tgccctaccc ctgcttcgtg gacgattctt tcatgatcat cctgaaccgc 2040

tgtttcccta tcttcatggt gctggcatgg atctactctg tctccatgac tgtgaagagc 2100

atcgtcttgg agaaggagtt gcgactgaag gagaccttga aaaatcaggg tgtctccaat 2160

gcagtgattt ggtgtacctg gttcctggac agcttctcca tcatgtcgat gagcatcttc 2220

ctcctgacga tattcatcat gcatggaaga atcctacatt acagcgaccc attcatcctc 2280

ttcctgttct tgttggcttt ctccactgcc accatcatgc tgtgctttct gctcagcacc 2340

ttcttctcca aggccagtct ggcagcagcc tgtagtggtg tcatctattt caccctctac 2400

ctgccacaca tcctgtgctt cgcctggcag gaccgcatga ccgctgagct gaagaaggct 2460

gtgagcttac tgtctccggt ggcatttgga tttggcactg agtacctggt tcgctttgaa 2520

gagcaaggcc tggggctgca gtggagcaac atcgggaaca gtcccacgga aggggacgaa 2580

ttcagcttcc tgctgtccat gcagatgatg ctccttgatg ctgcgtgcta tggcttactc 2640

gcttggtacc ttgatcaggt gtttccagga gactatggaa ccccacttcc ttggtacttt 2700

cttctacaag agtcgtattg gcttagcggt gaagggtgtt caaccagaga agaaagagcc 2760

ctggaaaaga ccgagcccct aacagaggaa acggaggatc cagagcaccc agaaggaata 2820

cacgactcct tctttgaacg tgagcatcca gggtgggttc ctggggtatg cgtgaagaat 2880

ctggtaaaga tttttgagcc ctgtggccgg ccagctgtgg accgtctgaa catcaccttc 2940

tacgagaacc agatcaccgc attcctgggc cacaatggag ctgggaaaac caccaccttg 3000

tccatcctga cgggtctgtt gccaccaacc tctgggactg tgctcgttgg gggaagggac 3060

attgaaacca gcctggatgc agtccggcag agccttggca tgtgtccaca gcacaacatc 3120

ctgttccacc acctcacggt ggctgagcac atgctgttct atgcccagct gaaaggaaag 3180

tcccaggagg aggcccagct ggagatggaa gccatgttgg aggacacagg cctccaccac 3240

aagcggaatg aagaggctca ggacctatca ggtggcatgc agagaaagct gtcggttgcc 3300

attgcctttg tgggagatgc caaggtggtg attctggacg aacccacctc tggggtggac 3360

ccttactcga gacgctcaat ctgggatctg ctcctgaagt atcgctcagg cagaaccatc 3420

atcatgccca ctcaccacat ggacgaggcc gaccaccaag gggaccgcat tgccatcatt 3480

gcccagggaa ggctctactg ctcaggcacc ccactcttcc tgaagaactg ctttggcaca 3540

ggcttgtact taaccttggt gcgcaagatg aaaaacatcc agagccaaag gaaaggcagt 3600

gaggggacct gcagctgctc gtctaagggt ttctccacca cgtgtccagc ccacgtcgat 3660

gacctaactc cagaacaagt cctggatggg gatgtaaatg agctgatgga tgtagttctc 3720

caccatgttc cagaggcaaa gctggtggag tgcattggtc aagaacttat cttccttctt 3780

ccaaataaga acttcaagca cagagcatat gccagccttt tcagagagct ggaggagacg 3840

ctggctgacc ttggtctcag cagttttgga atttctgaca ctcccctgga agagattttt 3900

ctgaaggtca cggaggattc tgattcagga cctctgtttg cgggtggcgc tcagcagaaa 3960

agagaaaacg tcaacccccg acacccctgc ttgggtccca gagagaaggc tggacagaca 4020

ccccaggact ccaatgtctg ctccccaggg gcgccggctg ctcacccaga gggccagcct 4080

cccccagagc cagagtgccc aggcccgcag ctcaacacgg ggacacagct ggtcctccag 4140

catgtgcagg cgctgctggt caagagattc caacacacca tccgcagcca caaggacttc 4200

ctggcgcaga tcgtgctccc ggctaccttt gtgtttttgg ctctgatgct ttctattgtt 4260

atccttcctt ttggcgaata ccccgctttg acccttcacc cctggatata tgggcagcag 4320

tacaccttct tcagcatgga tgaaccaggc agtgagcagt tcacggtact tgcagacgtc 4380

ctcctgaata agccaggctt tggcaaccgc tgcctgaagg aagggtggct tccggagtac 4440

ccctgtggca actcaacacc ctggaagact ccttctgtgt ccccaaacat cacccagctg 4500

ttccagaagc agaaatggac acaggtcaac ccttcaccat cctgcaggtg cagcaccagg 4560

gagaagctca ccatgctgcc agagtgcccc gagggtgccg ggggcctccc gccccccag 4620

agaacacagc gcagcacgga aattctacaa gacctgacgg acaggaacat ctccgacttc 4680

ttggtaaaaa cgtatcctgc tcttataaga agcagcttaa agagcaaatt ctgggtcaat 4740

gaacagaggt atggaggaat ttccattgga ggaaagctcc cagtcgtccc catcacgggg 4800

gaagcacttg ttgggttttt aagcgacctt ggccggatca tgaatgtgag cgggggccct 4860

atcactagag aggcctctaa agaaatacct gatttcctta aacatctaga aactgaagac 4920

aacattaagg tgtggtttaa taacaaaggc tggcatgccc tggtcagctt tctcaatgtg 4980

gcccacaacg ccatcttacg ggccagcctg cctaaggaca ggagccccga ggagtatgga 5040

atcaccgtca ttagccaacc cctgaacctg accaaggagc agctctcaga gattacagtg 5100

ctgaccactt cagtggatgc tgtggttgcc atctgcgtga ttttctccat gtccttcgtc 5160

ccagccagct ttgtccttta tttgatccag gagcgggtga acaaatccaa gcacctccag 5220

tttatcagtg gagtgagccc caccacctac tgggtgacca acttcctctg ggacatcatg 5280

aattattccg tgagtgctgg gctggtggtg ggcatcttca tcgggtttca gaagaaagcc 5340

tacacttctc cagaaaacct tcctgccctt gtggcactgc tcctgctgta tggatgggcg 5400

gtcattccca tgatgtaccc agcatccttc ctgtttgatg tccccagcac agcctatgtg 5460

gctttatctt gtgctaatct gttcatcggc atcaacagca gtgctattac cttcatcttg 5520
gaattatttg ataataaccg gacgctgctc aggttcaacg ccgtgctgag gaagctgctc 5580
attgtcttcc cccacttctg cctgggccgg ggcctcattg accttgcact gagccaggct 5640
gtgacagatg tctatgcccg gtttggtgag gagcactctg caaatccgtt ccactgggac 5700
ctgattggga agaacctgtt tgccatggtg gtggaagggg tggtgtactt cctcctgacc 5760
ctgctggtcc agcgccactt cttcctctcc caatggattg ccgagcccac taaggagccc 5820
attgttgatg aagatgatga tgtggctgaa gaaagacaaa gaattattac tggtggaaat 5880
aaaactgaca tcttaaggct acatgaacta accaagattt atctgggcac ctccagccca 5940
gcagtggaca ggctgtgtgt cggagttcgc cctggagagt gctttggcct cctgggagtg 6000
aatggtgccg gcaaaacaac cacattcaag atgctcactg gggacaccac agtgacctca 6060
ggggatgcca ccgtagcagg caagagtatt ttaaccaata tttctgaagt ccatcaaaat 6120
atgggctact gtcctcagtt tgatgcaatc gatgagctgc tcacaggacg agaacatctt 6180
tacctttatg cccggcttcg aggtgtacca gcagaagaaa tcgaaaaggt tgcaaactgg 6240
agtattaaga gcctgggcct gactgtctac gccgactgcc tggctggcac gtacagtggg 6300
ggcaacaagc ggaaactctc cacagccatc gcactcattg gctgcccacc gctggtgctg 6360
ctggatgagc ccaccacagg gatggacccc caggcacgcc gcatgctgtg gaacgtcatc 6420
gtgagcatca tcagaaaagg gagggctgtg gtcctcacat cccacagcat ggaagaatgt 6480
gaggcactgt gtacccggct ggccatcatg gtaaagggcg cctttcgatg tatgggcacc 6540
attcagcatc tcaagtccaa atttggagat ggctatatcg tcacaatgaa gatcaaatcc 6600
ccgaaggacg acctgcttcc tgacctgaac cctgtggagc agttcttcca ggggaacttc 6660
ccaggcagtg tgcagaggga gaggcactac aacatgctcc agttccaggt ctcctcctcc 6720
tccctggcga ggatcttcca gctcctcctc tcccacaagg acagcctgct catcgaggag 6780
tactcagtca cacagaccac actggaccag gtgtttgtaa attttgctaa acagcagact 6840
gaaagtcatg acctccctct gcaccctcga gctgctggag ccagtcgaca agcccaggac 6900
tgatctttca caccgctcgt tcctgcagcc agaaaggaac tctgggcagc tggaggcgca 6960
ggagcctgtg cccatatggt catccaaatg gactggccca gcgtaaatga ccccactgca 7020
gcagaaaaca aacacacgag gagcatgcag cgaattcaga aagaggtctt tcagaaggaa 7080
accgaaactg acttgctcac ctggaacacc tgatggtgaa accaaacaaa tacaaaatcc 7140
ttctccagac cccagaacta gaaaccccgg gccatcccac tagcagcttt ggcctccata 7200
ttgctctcat ttcaagcaga tctgcttttc tgcatgtttg tctgtgtgtc tgcgttgtgt 7260

gtgattttca tggaaaaata aaatgcaaat gcactcatca caaaaaaaaa aaaaaaaa 7318

<210> 13
<211> 2663
<212> DNA
<213> Homo Sapiens

<400> 13
ggcacgaggc tggtgtttag caactccgac cacctgcctg ctgaggggct agagccctca 60

gcccagaccc tgtgcccccg gccgggctct catgcgtgga atggtgctgt gccccttgcc 120

agcaggccag gctcaccatg gtgccgcatg ccatcttggc acgggggagg gacgtgtgca 180

ggcggaatgg actcctcatc ctgtctgtgc tgtctgtcat cgtgggctgc ctcctcggct 240

tcttcttgag gacccggcgc ctctcaccac aggaaattag ttacttccag ttccctggag 300

agctcctgat gaggatgctg aagatgatga tcctgccact ggtggtctcc agcttgatgt 360

ccggacttgc ctccctggat gccaagacct ctagccgcct gggcgtcctc accgtggcgt 420

actacctgtg gaccaccttc atggctgtca tcgtgggcat cttcatggtc tccatcatcc 480

acccaggcag cgcggcccag aaggagacca cggagcagag tgggaagccc atcatgagct 540

cagccgatgc cctgttggac ctcatccgga acatgttccc agccaaccta gtagaagcca 600

cattcaaaca gtaccgcacc aagaccaccc cagttgtcaa gtcccccaag gtggcaccag 660

aggaggcccc tcctcggcgg atcctcatct acggggtcca ggaggagaat ggctcccatg 720

tgcagaactt cgccctggac ctgaccccgc cgcccgaggt cgtttacaag tcagagccgg 780

gcaccagcga tggcatgaat gtgctgggca tcgtcttctt ctctgccacc atgggcatca 840

tgctgggccg catgggtgac agcggggccc ccctggtcag cttctgccag tgcctcaatg 900

agtcggtcat gaagatcgtg gcggtggctg tgtggtattt ccccttcggc attgtgttcc 960

tcattgcggg taagatcctg gagatggacg accccagggc cgtcggcaag aagctgggct 1020

tctactcagt caccgtggtg tgcgggctgg tgctccacgg gctctttatc ctgccctgc 1080

tctacttctt catcaccaag aagaatccca tcgtcttcat ccgcggcatc ctgcaggctc 1140

tgctcatcgc gctggccacc tcctccagct cagccacact gcccatcacc ttcaagtgcc 1200

tgctggagaa caaccacatc gaccggcgca tcgctcgctt cgtgctgccc gtgggtgcca 1260

ccatcaacat ggacggcact gcgctctacg aggctgtggc cgccatcttc atcgcccagg 1320

tcaacaacta cgagctggac tttggccaga tcatcaccat cagtatcaca gccactgcag 1380

ccagcattgg ggcagctggc atcccccagg ccggcctcgt caccatggtc atcgtgctca 1440

cctccgtggg actgcccacc gatgacatca ccctcatcat tgccgttgac tgggctctgg 1500

accgtttccg caccatgatt aacgtgctgg gtgatgcgct ggcagcgggg atcatggccc 1560

atatatgtcg gaaggatttt gcccgggaca caggcaccga gaaactgctg ccctgcgaga 1620

ccaagccagt gagcctccag gagatcgtgg cagcccagca gaatggctgt gtgaagagtg 1680

tagccgaggc ctccgagctc accctgggcc ccacctgccc ccaccacgtc cccgttcaag 1740

tggagcggga tgaggagctg cccgctgcga gtctgaacca ctgcaccatc cagatcagcg 1800

agctggagac caatgtctga gcctgcggag ctgcaggggc aggcgaggcc tccaggggca 1860

gggtcctgag gcaggaactc gactctccaa ccctcctgag cagccggcag gggccaggat 1920

cacacattct tctcaccctt gagaggctgg aattaacccc gcttgacgga aaatgtatct 1980

cagagaaggg aaaggctgca tgggggagcc ccatccaggg agtgatgggc ccggcattgc 2040

ctgaggcccc gctgtgacag tttccccggt gtgagcccgg tgagggcggc aggcaggggt 2100

tatccggccc cactttctgg atgacagact tgaggctctg agagctgaaa acacttgtcc 2160

aaggtctcac gttaaggtca agacactaac tcaaatcttt caagccccgc ctctcctctt 2220

ggaggacagg gcagcctgca gctgtgtcca ggcccaggcc ccaccccata acaggtggcc 2280

tcagccacac agttctcccc aaggggagca gcccagggcc aagccccgct gccttcccca 2340

ggccacagtg cgtccagtct cctgtcctgc cacgtgtctt ttgcaaagct ccttggatgt 2400

ggagacagat gtctttacta gagctgaaag gccccccttga cacatccagg ccaacctccc 2460

atggaatagg taggcaagcc aggactccgg gaaggaggtg cagccaggat gctctggtgg 2520

agctgccgat ggggccctgg tgtcagaact ccccaaaggc ctgtgcgtcc aagtggagtc 2580

aggttttcta ttcctttctg tgtttgcaaa ttcagtgtta actaaataaa ggtattttgt 2640

ttttcaaaaa aaaaaaaaaa aaa 2663

<210> 14
<211> 2993
<212> DNA
<213> Homo Sapiens

<400> 14
gggcctgcag ttggcagaag ggtcccgggc ccagagccag cggggccgtg ctgagacggc 60

gtacgtgccc tgcgtgagtg cgtggcggcg gcgcgtgcgc taggggagtg ggcggtgagg 120

cctggtccac gtgcgtccct tcccgggacc cccgcagctt ggcgcccagc ggctacgtga 180

gccaaggcac ccggatgtcc gcgcccctct ccgagtgaca agtcccggcc tccggtcccg 240

cagtgcccgc agcctcggcc ggcgtccacg cattgccatg gtgactgtgg gcaactactg 300

cgaggccgaa gggcccgtgg gtccggcctg gatgcaggat ggcctgagtc cctgcttctt 360

cttcacgctc gtgccctcga cgcggatggc tctagggact ctggccttgg tgctggctct 420

tccctgcaga cgccgggagc ggcccgctgg tgctgattcg ctgtcttggg gggccggccc 480

tcgcatctct ccctacgtgc tgcagctgct tctggccaca cttcaggcgg cgctgcccct 540

ggccggcctg gctggccggg tgggcactgc ccgggggccc ccactgccaa gctatctact 600

tctggcctcc gtgctggaga gtctggccgg cgcctgtggc ctgtggctgc ttgtcgtgga 660

gcggagccag gcacggcagc gtctggcaat gggcatctgg atcaagttca ggcacagccc 720

tggtctcctg ctcctctgga ctgtggcgtt tgcagctgag aacttggccc tggtgtcttg 780

gaacagccca cagtggtggt gggcaagggc agacttgggc caacaggttc agtttagcct 840

gtgggtgctg cggtatgtgg tctctggagg gctgtttgtc ctgggtctct gggcccctgg 900

acttcgtccc cagtcctata cattgcaggt tcatgaagag gaccaagatg tggaaaggag 960

ccaggttcgg tcagcagccc aacagtctac ctggcgagat tttggcagga agctccgcct 1020

cctgagtggc tacctgtggc ctcgagggag tccagctctg cagctggtgg tgctcatctg 1080

cctggggctc atgggtttgg aacgggcact caatgtgttg gtgcctatat tctataggaa 1140

cattgtgaac ttgctgactg agaaggcacc ttggaactct ctggcctgga ctgttaccag 1200

ttacgtcttc ctcaagttcc tccaggggggg tggcactggc agtacaggct tcgtgagcaa 1260

cctgcgcacc ttcctgtgga tccgggtgca gcagttcacg tctcggcggg tggagctgct 1320

catcttctcc cacctgcacg agctctcact gcgctggcac ctggggcgcc gcacagggga 1380

ggtgctgcgg atcgcggatc ggggcacatc cagtgtcaca gggctgctca gctacctggt 1440

gttcaatgtc atccccacgc tggccgacat catcattggc atcatctact tcagcatgtt 1500

cttcaacgcc tggtttggcc tcattgtgtt cctgtgcatg agtctttacc tcaccctgac 1560

cattgtggtc actgagtgga gaaccaagtt tcgtcgtgct atgaacacac aggagaacgc 1620

tacccgggca cgagcagtgg actctctgct aaacttcgag acggtgaagt attacaacgc 1680

cgagagttac gaagtggaac gctatcgaga ggccatcatc aaatatcagg gtttggagtg 1740

gaagtcgagc gcttcactgg ttttactaaa tcagacccag aacctggtga ttgggctcgg 1800

gctcctcgcc ggctccctgc tttgcgcata ctttgtcact gagcagaagc tacaggttgg 1860

ggactatgtg ctctttggca cctacattat ccagctgtac atgcccctca attggtttgg 1920

cacctactac aggatgatcc agaccaactt cattgacatg gagaacatgt ttgacttgct 1980

gaaagaggag acagaagtga aggaccttcc tggagcaggg ccccttcgct ttcagaaggg 2040

ccgtattgag tttgagaacg tgcacttcag ctatgccgat gggcgggaga ctctgcagga 2100

cgtgtctttc actgtgatgc ctggacagac acttgccctg gtgggcccat ctggggcagg 2160

gaagagcaca attttgcgcc tgctgtttcg cttctacgac atcagctctg gctgcatccg 2220

aatagatggg caggacattt cacaggtgac ccaggcctct ctccggtctc acattggagt 2280

tgtgccccaa gacactgtcc tctttaatga caccatcgcc gacaatatcc gttacggccg 2340

tgtcacagct gggaatgatg aggtggaggc tgctgctcag gctgcaggca tccatgatgc 2400

cattatggct ttccctgaag ggtacaggac acaggtgggc gagcggggac tgaagctgag 2460

cggcggggag aagcagcgcg tcgccattgc ccgcaccatc ctcaaggctc cgggcatcat 2520

tctgctggat gaggcaacgt cagcgctgga tacatctaat gagagggcca tccaggcttc 2580

tctggccaaa gtctgtgcca accgcaccac catcgtagtg gcacacaggc tctcaactgt 2640

ggtcaatgct gaccagatcc tcgtcatcaa ggatggctgc atcgtggaga ggggacgaca 2700

cgaggctctg ttgtcccgag gtggggtgta tgctgacatg tggcagctgc agcagggaca 2760

ggaagaaacc tctgaagaca ctaagcctca gaccatggaa cggtgacaaa agtttggcca 2820

cttccctctc aaagactaac ccagaaggga ataagatgtg tctcctttcc ctggcttatt 2880

tcatcctggt cttggggtat ggtgctagct atggtaaggg aaagggacct ttccgaaaaa 2940

catcttttgg ggaaataaaa atgtggactg tgaaaaaaaa aaaaaaaaaa aaa 2993

<210> 15
<211> 1733
<212> DNA
<213> Homo Sapiens

<400> 15
gggcaatttg ttagttatcc gccgccacca agacgcggca cggcgcctgg accggagggg 60

ccccgcgcgg gcgcgaactt tgggctcggg cgagtgggtg gtgctccgcc cagcccgaga 120

cgggcgggcg cgcgggccaa tgggtgccgc ctcttggccg cgggggggcc cgacccgtgg 180

gtcccggcca ccagcgcccc agccccgagg ctcagaagcg gcaggcggag gcgcggtccg 240

ggcgctatgg ccatgcccgg cgggtctcac gcggctgccc ctcgcccggc gcgccttcgg 300

taggggcgc ccggggccca gctggcccgg ccatgctgct ggagacacag gacgcgctgt 360

acgtggcgct ggagctggtc atcgccgcgc tttcggtggc gggcaacgtg ctggtgtgcg 420

ccgcggtggg cacggcgaac actctgcaga cgcccaccaa ctacttcctg gtgtccctgg 480

ctgcggccga cgtggccgtg gggctcttcg ccatcccctt tgccatcacc atcagcctgg 540

gcttctgcac tgacttctac ggctgcctct tcctcgcctg cttcgtgctg gtgctcacgc 600

agagctccat cttcagcctt ctggccgtgg cagtcgacag atacctggcc atctgtgtcc 660

cgctcaggta taaaagtttg gtcacgggga cccgagcaag aggggtcatt gctgtcctct 720

gggtccttgc ctttggcatc ggattgactc cattcctggg gtggaacagt aaagacagtg 780

ccaccaacaa ctgcacagaa ccctgggatg gaaccacgaa tgaaagctgc tgccttgtga 840

agtgtctctt tgagaatgtg gtccccatga gctacatggt atatttcaat ttctttgggt 900

gtgttctgcc cccactgctt ataatgctgg tgatctacat taagatcttc ctggtggcct 960

gcaggcagct tcagcgcact gagctgatgg accactcgag gaccaccctc cagcgggaga 1020

tccatgcagc caagtcactg gccatgattg tggggatttt tgccctgtgc tggttacctg 1080

tgcatgctgt taactgtgtc actcttttcc agccagctca gggtaaaaat aagcccaagt 1140

gggcaatgaa tatggccatt cttctgtcac atgccaattc agttgtcaat cccattgtct 1200

atgcttaccg gaaccgagac ttccgctaca cttttcacaa aattatctcc aggtatcttc 1260

tctgccaagc agatgtcaag agtgggaatg gtcaggctgg ggtacagcct gctctcggtg 1320

tgggcctatg atctaggctc tcgcctcttc caggagaaga tacaaatcca caagaaacaa 1380

agaggacacg gctggttttc attgtgaaag atagctacac ctcacaagga aatggactgc 1440

ctctcttgag cacttccctg gagctaccac gtatctagct aatatgtatg tgtcagtagt 1500

aggctccaag gattgacaaa tatatttatg atctattcag ctgcttttac tgtgtggatt 1560

atgccaacag cttgaatgga ttctaacaga ctcttttgtt tttaaaagtc tgccttgttt 1620

atggtggaaa attactgaaa ctattttact gtgaaacagt gtgaactatt ataatgcaaa 1680

tacttttttaa cttagaggca atggaaaaat aaaagttgac tgtactaaaa atg 1733

<210> 16
<211> 3338
<212> DNA
<213> Homo Sapiens

<400> 16
cccagcccgg ccccgccgcc ccggctgcgc acgcgacgcc ccctccaggc cccgctcctg 60

cgccctattt ggtcattcgg ggggcaagcg gcgggagggg aaacgtgcgc ggccgaaggg 120

gaagcggagc cggcgccggc tgcgcagagg agccgctctc gccgccgcca cctcggctgg 180

gagcccacga ggctgccgca tcctgccctc ggaacaatgg gactcggcgc gcgaggtgct 240

tgggccgcgc tgctcctggg gacgctgcag gtgctagcgc tgctgggggc cgcccatgaa 300

agcgcagcca tggcggagac tctccaacat gtgccttctg accatacaaa tgaaacttcc 360

aacagtactg tgaaaccacc aacttcagtt gcctcagact ccagtaatac aacggtcacc 420

accatgaaac ctacagcggc atctaataca acaacaccag ggatggtctc aacaaatatg 480

acttctacca ccttaaagtc tacacccaaa acaacaagtg tttcacagaa cacatctcag 540

atatcaacat ccacaatgac cgtaacccac aatagttcag tgacatctgc tgcttcatca 600

gtaacaatca caacaactat gcattctgaa gcaaagaaag gatcaaaatt tgatactggg 660

agctttgttg gtggtattgt attaacgctg ggagttttat ctattcttta cattggatgc 720

aaaatgtatt actcaagaag aggcattcgg tatcgaacca tagatgaaca tgatgccatc 780

atttaaggaa atccatggac caaggatgga atacagattg atgctgccct atcaattaat 840

tttggtttat taatagttta aaacaatatt ctctttttga aaatagtata aacaggccat 900
gcatataatg tacagtgtat tacgtaaata tgtaaagatt cttcaaggta acaagggttt 960
gggttttgaa ataaacatct ggatcttata gaccgttcat acaatggttt tagcaagttc 1020
atagtaagac aaacaagtcc tatcttttt tttttggctg gggtgggggc attggtcaca 1080
tatgaccagt aattgaaaga cgtcatcact gaaagacaga atgccatctg ggcatacaaa 1140
taagaagttt gtcacagcac tcaggatttt gggtatcttt tgtagctcac ataaagaact 1200
tcagtgcttt tcagagctgg atatatctta attactaatg ccacacagaa attatacaat 1260
caaactagat ctgaagcata atttaagaaa aacatcaaca ttttttgtgc tttaaactgt 1320
agtagttggt ctagaaacaa aatactccaa gaaaaagaaa attttcaaat aaaacccaaa 1380
ataatagctt tgcttagccc tgttagggat ccattggagc attaaggagc acatattttt 1440
attaacttct tttgagcttt caatgttgat gtaatttttg ttctctgtgt aatttaggta 1500
aactgcagtg tttaacataa taatgtttta aagacttagt tgtcagtatt aaataatcct 1560
ggcattatag ggaaaaaacc tcctagaagt tagattattt gctactgtga gaatattgtc 1620
accactggaa gttactttag ttcatttaat tttaatttta tattttgtga atattttaag 1680
aactgtagag ctgctttcaa tatctagaaa tttttaattg agtgtaaaca cacctaactt 1740
taagaaaaag aaccgcttgt atgattttca aaagaacatt tagaattcta tagagtcaaa 1800
actatagcgt aatgctgtgt ttattaagcc agggattgtg ggacttcccc caggcaacta 1860
aacctgcagg atgaaaatgc tatattttct ttcatgcact gtcgatatta ctcagatttg 1920
gggaaatgac atttttatac taaaacaaac accaaaatat tttagaataa attcttagaa 1980
agttttgaga ggaattttta gagaggacat ttcctccttc ctgatttgga tattccctca 2040
aatccctcct cttactccat gctgaaggag aagtactctc agatgcatta tgttaatgga 2100
gagaaaaagc acagtattgt agagacacca atattagcta atgtattttg gagtgttttc 2160
cattttacag tttatattcc agcactcaaa actcagggtc aagttttaac aaaagaggta 2220
tgtagtcaca gtaaatacta agatggcatt tctatctcag agggccaaag tgaatcacac 2280
cagtttctga aggtcctaaa aatagctcag atgtcctaat gaacatgcac ctacatttaa 2340
taggagtaca ataaaactgt tgtcagcttt tgttttacag agaacgctag atattaagaa 2400
ttttgaaatg gatcatttct acttgctgtg cattttaacc aataatctga tgaatataga 2460
aaaaaatgat ccaaaatatg gatatgattg gatgtatgta acacatacat ggagtatgga 2520
ggaaattttc tgaaaaatac atttagatta gtttagtttg aaggagaggt gggctgatgg 2580
ctgagttgta tgttactaac ttggccctga ctggttgtgc aaccattgct tcatttcttt 2640

gcaaaatgta gttaagatat actttattct aatgaaggcc ttttaaattt gtccactgca 2700

ttcttggtat ttcactactt caagtcagtc agaacttcgt agaccgacct gaagtttctt 2760

tttgaatact tgtttcttta gcactttgaa gatagaaaaa ccacttttta agtactaagt 2820

catcatttgc cttgaaagtt tcctctgcat tgggtttgaa gtagtttagt tatgtctttt 2880

tctctgtatg taagtagtat aatttgttac tttcaaatac ccgtactttg aatgtaggtt 2940

tttttgttgt tgttatctat aaaaattgag ggaaatggtt atgcaaaaaa atattttgct 3000

ttggaccata tttcttaagc ataaaaaaat gctcagtttt gcttgcattc cttgagaatg 3060

tatttatctg aagatcaaaa caaacaatcc agatgtataa gtactaggca gaagccaatt 3120

ttaaaatttc cttgaataat ccatgaaagg aataattcaa atacagataa acagagttgg 3180

cagtatatta tagtgataat tttgtatttt caamaaaaaa aaagttaaac tcttcttttc 3240

tttttattat aatgaccagc ttttggtatt tcattgttac caagttctat ttttagataa 3300

aattgttctc cttctaaaaa aaaaaaaaaa aaaaaaaa 3338

<210> 17
<211> 1214
<212> DNA
<213> Homo Sapiens

<400> 17
gtctacaccc cctcctcaca cgcacttcac ctgggtcggg attctcaggt catgaacggt 60

cccagccacc tccgggcagg gcgggtgagg acggggacgg ggcgtgtcca actggctgtg 120

ggctcttgaa acccgagcat ggcacagcac ggggcgatgg gcgcgtttcg ggccctgtgc 180

ggcctggcgc tgctgtgcgc gctcagcctg ggtcagcgcc ccaccggggg tcccgggtgc 240

ggccctgggc gcctcctgct tgggacggga acggacgcgc gctgctgccg ggttcacacg 300

acgcgctgct gccgcgatta cccgggcgag gagtgctgtt ccgagtggga ctgcatgtgt 360

gtccagcctg aattccactg cggagaccct tgctgcacga cctgccggca ccacccttgt 420

cccccaggcc agggggtaca gtcccagggg aaattcagtt ttggcttcca gtgtatcgac 480

tgtgcctcgg ggaccttctc cgggggccac gaaggccact gcaaaccttg gacagactgc 540

acccagttcg ggtttctcac tgtgttccct gggaacaaga cccacaacgc tgtgtgcgtc 600

ccagggtccc cgccggcaga gccgcttggg tggctgaccg tcgtcctcct ggccgtggcc 660

gcctgcgtcc tcctcctgac ctcggcccag cttggactgc acatctggca gctgaggagt 720

cagtgcatgt ggccccgaga gacccagctg ctgctggagg tgccgccgtc gaccgaagac 780

gccagaagct gccagttccc cgaggaagag cggggcgagc gatcggcaga ggagaagggg 840

cggctgggag acctgtgggt gtgagcctgg ccgtcctccg gggccaccga ccgcagccag 900

ccctcccca ggagctcccc aggccgcagg ggctctgcgt tctgctctgg gccgggccct 960

gctcccctgg cagcagaagt gggtgcagga aggtggcagt gaccagcgcc ctggaccatg 1020

cagttcggcg gccgcggctg ggccctgcag gagggagaga gagacacagt catggccccc 1080

ttcctccctt gctggccctg atggggtggg gtcttaggac gggaggctgt gtccgtgggt 1140

gtgcagtgcc cagcacggga cccggctgca ggggaccttc aataaacact tgtccagtga 1200

aaaaaaaaaa aaaa 1214

<210> 18
<211> 2322
<212> DNA
<213> Homo Sapiens

<400> 18
agatccgcga gcccgtcagc ctgcgccatg ggctgcgacg gccgcgtgtc ggggctgctc 60

cgccgcaacc tgcagcccac gctcacctac tggagcgtct tcttcagctt cggcctgtgc 120

atcgccttcc tggggcccac gctgctggac ctgcgctgtc agacgcacag ctcgctgccc 180

cagatctcct gggtcttctt ctcgcagcag ctctgcctcc tgctgggcag cgccctcggg 240

ggcgtcttca aaaggaccct ggcccagtca ctatgggccc tgttcacctc ctctctggcc 300

atctccctgg tgtttgccgt catccccttc tgccgcgacg tgaaggtgct ggcctcagtc 360

atggcgctgg cgggcttggc catgggctgc atcgacaccg tggccaacat gcagctggta 420

aggatgtacc agaaggactc ggccgtcttc ctccaggtgc tccatttctt cgtgggcttt 480

ggtgctctgc tgagcccccc tattgctgac cctttcctgt ctgaggccaa ctgcttgcct 540

gccaatagca cggccaacac cacctcccga ggccacctgt tccatgtctc cagggtgctg 600

ggccagcacc acgtagatgc caagccttgg tccaaccaga cgttcccagg gctgactcca 660

aaggacgggg cagggacccg agtgtcctat gccttctgga tcatggccct catcgatctt 720

ccagtgccca tggctgtgct gatgctgctg tccaaggagc ggctgctgac ctgctgtccc 780

cagaggaggc ccctgcttct gtctgctgat gagcttgcct tggagacaca gcctcctgag 840

aaggaagatg cctcctcact gcccccaaag tttcagtcac acctagggca tgaggacctg 900

ttcagctgct gccaaaggaa gaacctcaga ggagcccctt attccttctt tgccatccac 960

atcacgggcg ccctggtact gttcatgacg gatgggttga cgggtgccta ttccgccttc 1020

gtgtacagct atgctgtgga gaagcccctg tctgtgggac acaaggtggc tggctacctc 1080

cccagcctct tctggggctt catcacactg ggccggctcc tctccattcc catatcctca 1140

agaatgaagc cggccaccat ggttttcatc aacgtggttg gcgtggtggt gacgttcctg 1200

gtgctgctta ttttctccta caacgtcgtc ttcctgttcg tggggacggc aagcctgggc 1260

ctgtttctca gcagcacctt ccccagcatg ctggcctaca cggaggactc gctgcagtac 1320

aaaggctgtg caaccacagt gctggtgaca ggggcaggag ttggcgagat ggtgctgcag 1380

atgctggttg gttcgatatt ccaggctcag ggcagctata gtttcctggt ctgtggcgtg 1440

atctttggtt gtctggcttt taccttctat atcttgctcc tgtttttcca caggatgcac 1500

cctggactcc catcagttcc tacccaagac agatcaattg gaatggaaaa ctctgagtgc 1560

taccagaggt aaaactgggt gaagaaggca agagaagact ttcagcctct tgatcaccag 1620

cacgaccata ctgtttcaga aagctgggtg gtggtggagg cgctctctca atggctattc 1680

aagtcttctc cactaaaact tggttgggta gaggaaatta aattgagtcc tggtacctgg 1740

tcaaaatcat tagaagttta cctggcttct caagttatct tcttccctgg ttcagactgt 1800

tggtaagagc tgtccagata cccagatggg aaggaaggag acagccgcgc gcttcactcc 1860

atttgtcacc tcatgcatgg accatactct gggtttgaga tcattcttca ttgaagtttg 1920

taaaaatagg ttgaaattgt aaagctccat gatcactgct atatgtagat atatttcaat 1980

ttaagcaaaa caagctgcaa gttattccct ggcatgctca aaggattttc gtgcttttca 2040

cttaatagtc caaagtctct taaattcctg ctgcagatat caatagctta tctatattct 2100

caaacaccaa aaggaaaagt tgaatcttgc tctctttggt atactaatgt agtggtatgc 2160

taagctggct cataccaact tagaaaagct gattgtaaaa ttttcatttt gacagctggt 2220

tattaaatgc agccattatt aaaaatcaaa tcatacaaac ttataattaa atcaattaca 2280

tttaaaacaa aggtaataaa tattcaaagc atatcacttc ct 2322

<210> 19
<211> 5361
<212> DNA
<213> *Homo Sapiens*

<400> 19
ctgcaaaccc agcgcaacta cggtcccccg gtcagaccca ggatggggcc agaacggaca 60

ggggccgcgc cgctgccgct gctgctggtg ttagcgctca gtcaaggcat tttaaattgt 120

tgtttggcct acaatgttgg tctcccagaa gcaaaaatat tttccggtcc ttcaagtgaa 180

cagtttgggt atgcagtgca gcagtttata aatccaaaag gcaactggtt actggttggt 240

tcaccctgga gtggctttcc tgagaaccga atgggagatg tgtataaatg tcctgttgac 300

ctatccactg ccacatgtga aaaactaaat ttgcaaactt caacaagcat tccaaatgtt 360

actgagatga aaaccaacat gagcctcggc ttgatcctca ccaggaacat gggaactgga 420

ggttttctca catgtggtcc tctgtgggca cagcaatgtg ggaatcagta ttacacaacg 480

ggtgtgtgtt ctgacatcag tcctgatttt cagctctcag ccagcttctc acctgcaact 540

cagccctgcc cttccctcat agatgttgtg gttgtgtgtg atgaatcaaa tagtatttat 600
ccttgggatg cagtaaagaa ttttttggaa aaatttgtac aaggccttga tataggcccc 660
acaaagacac aggtggggtt aattcagtat gccaataatc caagagttgt gtttaacttg 720
aacacatata aaaccaaaga agaaatgatt gtagcaacat cccagacatc ccaatatggt 780
ggggacctca caaacacatt cggagcaatt caatatgcaa gaaaatatgc ctattcagca 840
gcttctggtg ggcgacgaag tgctacgaaa gtaatggtag ttgtaactga cggtgaatca 900
catgatggtt caatgttgaa agctgtgatt gatcaatgca accatgacaa tatactgagg 960
tttggcatag cagttcttgg gtacttaaac agaaacgccc ttgatactaa aaatttaata 1020
aaagaaataa aagcgatcgc tagtattcca acagaaagat acttttttcaa tgtgtctgat 1080
gaagcagctc tactagaaaa ggctgggaca ttaggagaac aaatttttcag cattgaaggt 1140
actgttcaag gaggagacaa ctttcagatg gaaatgtcac aagtgggatt cagtgcagat 1200
tactcttctc aaaatgatat tctgatgctg ggtgcagtgg gagcttttgg ctggagtggg 1260
accattgtcc agaagacatc tcatggccat ttgatctttc ctaaacaagc ctttgaccaa 1320
attctgcagg acagaaatca cagttcatat ttaggttact ctgtggctgc aatttctact 1380
ggagaaagca ctcactttgt tgctggtgct cctcgggcaa attataccgg ccagatagtg 1440
ctatatagtg tgaatgagaa tggcaatatc acggttattc aggctcaccg aggtgaccag 1500
attggctcct attttggtag tgtgctgtgt tcagttgatg tggataaaga caccattaca 1560
gacgtgctct tggtaggtgc accaatgtac atgagtgacc taaagaaaga ggaaggaaga 1620
gtctacctgt ttactatcaa aaagggcatt ttgggtcagc accaatttct tgaaggcccc 1680
gagggcattg aaaacactcg atttggttca gcaattgcag ctctttcaga catcaacatg 1740
gatggcttta atgatgtgat tgttggttca ccactagaaa atcagaattc tggagctgta 1800
tacatttaca atggtcatca gggcactatc cgcacaaagt attcccagaa aatcttggga 1860
tccgatggag cctttaggag ccatctccag tactttggga ggtccttgga tggctatgga 1920
gatttaaatg gggattccat caccgatgtg tctattggtg cctttggaca agtggttcaa 1980
ctctggtcac aaagtattgc tgatgtagct atagaagctt cattcacacc agaaaaaatc 2040
actttggtca acaagaatgc tcagataatt ctcaaactct gcttcagtgc aaagttcaga 2100
cctactaagc aaaacaatca agtggccatt gtatataaca tcacacttga tgcagatgga 2160
ttttcatcca gagtaacctc caggggggtta tttaaagaaa acaatgaaag gtgcctgcag 2220
aagaatatgg tagtaaatca agcacagagt tgccccgagc acatcattta tatacaggag 2280
ccctctgatg ttgtcaactc tttggatttg cgtgtggaca tcagtctgga aaaccctggc 2340
actagccctg cccttgaagc ctattctgag actgccaagg tcttcagtat tcctttccac 2400

aaagactgtg gtgaggatgg actttgcatt tctgatctag tcctagatgt ccgacaaata 2460
ccagctgctc aagaacaacc ctttattgtc agcaaccaaa acaaaaggtt aacattttca 2520
gtaacactga aaaataaaag ggaaagtgca tacaacactg gaattgttgt tgatttttca 2580
gaaaacttgt tttttgcatc attctcccta ccggttgatg ggacagaagt aacatgccag 2640
gtggctgcat ctcagaagtc tgttgcctgc gatgtaggct accctgcttt aaagagagaa 2700
caacaggtga cttttactat taactttgac ttcaatcttc aaaaccttca gaatcaggcg 2760
tctctcagtt tccaagcctt aagtgaaagc caagaagaaa acaaggctga taatttggtc 2820
aacctcaaaa ttcctctcct gtatgatgct gaaattcact taacaagatc taccaacata 2880
aatttttatg aaatctcttc ggatgggaat gttccttcaa tcgtgcacag ttttgaagat 2940
gttggtccaa aattcatctt ctccctgaag gtaacaacag gaagtgttcc agtaagcatg 3000
gcaactgtaa tcatccacat ccctcagtat accaaagaaa agaacccact gatgtaccta 3060
actggggtgc aaacagacaa ggctggtgac atcagttgta atgcagatat caatccactg 3120
aaaataggac aaacatcttc ttctgtatct ttcaaaagtg aaaatttcag gcacaccaaa 3180
gaattgaact gcagaactgc ttcctgtagt aatgttacct gctggttgaa agacgttcac 3240
atgaaaggag aatactttgt taatgtgact accagaattt ggaacgggac tttcgcatca 3300
tcaacgttcc agacagtaca gctaacggca gctgcagaaa tcaacaccta taaccctgag 3360
atatatgtga ttgaagataa cactgttacg attcccctga tgataatgaa acctgatgag 3420
aaagccgaag taccaacagg agttataata ggaagtataa ttgctggaat ccttttgctg 3480
ttagctctgg ttgcaatttt atggaagctc ggcttcttca aaagaaaata tgaaaagatg 3540
accaaaaatc cagatgagat tgatgagacc acagagctca gtagctgaac cagcagacct 3600
acctgcagtg ggaaccggca gcatcccagc cagggtttgc tgtttgcgtg catggatttc 3660
tttttaaatc ccatattttt tttatcatgt cgtaggtaaa ctaacctggt attttaagag 3720
aaaactgcag gtcagtttgg atgaagaaat tgtgggggggt gggggaggtg cgggggggcag 3780
gtagggaaat aatagggaaa atacctattt tatatgatgg gggaaaaaaa gtaatcttta 3840
aactggctgg cccagagttt acattctaat ttgcattgtg tcagaaacat gaaatgcttc 3900
caagcatgac aacttttaaa gaaaaatatg atactctcag attttaaggg ggaaaactgt 3960
tctctttaaa atatttgtct ttaaacagca actacagaag tggaagtgct tgatatgtaa 4020
gtacttccac ttgtgtatat tttaatgaat attgatgtta acaagagggg aaaacaaaac 4080
acaggttttt tcaatttatg ctgctcatcc aaagttgcca cagatgatac ttccaagtga 4140
taattttatt tataaactag gtaaaatttg ttgttggttc cttttatacc acggctgccc 4200

cttccacacc ccatcttgct ctaatgatca aaacatgctt gaataactga gcttagagta 4260

tacctcctat atgtccattt aagttaggag agggggcgat atagagacta aggcacaaaa 4320

ttttgtttaa aactcagaat ataacattta tgtaaaatcc catctgctag aagcccatcc 4380

tgtgccagag gaaggaaaag gaggaaattt cctttctctt ttaggaggca caacagttct 4440

cttctaggat ttgtttggct gactggcagt aacctagtga atttttgaaa gatgagtaat 4500

ttctttggca accttcctcc tcccttactg aaccactctc ccacctcctg gtggtaccat 4560

tattatagaa gccctctaca gcctgacttt ctctccagcg gtccaaagtt atcccctcct 4620

ttacccctca tccaaagttc ccactccttc aggacagctg ctgtgcatta gatattaggg 4680

gggaaagtca tctgtttaat ttacacactt gcatgaatta ctgtatataa actccttaac 4740

ttcagggagc tattttcatt tagtgctaaa caagtaagaa aaataagcta gagtgaattt 4800

ctaaatgttg gaatgttatg ggatgtaaac aatgtaaagt aaaacactct caggatttca 4860

ccagaagtta cagatgaggc actggaaacc accaccaaat tagcaggtgc accttctgtg 4920

gctgtcttgt ttctgaagta ctttttcttc cacaagagtg aatttgacct aggcaagttt 4980

gttcaaaagg tagatcctga gatgatttgg tcagattggg ataaggccca gcaatctgca 5040

ttttaacaag caccccagtc actaggatgc agatggacca cactttgaga aacaccaccc 5100

atttctactt tttgcacctt attttctctg ttcctgagcc cccacattct ctaggagaaa 5160

cttagattaa aattcacaga cactacatat ctaaagcttt gacaagtcct tgacctctat 5220

aaacttcaga gtcctcatta taaaatggga agactgagct ggagttcagc agtgatgctt 5280

tttagtttta aaagtctatg atctgatctg gacttcctat aatacaaata cacaatcctc 5340

caagaatttg acttggaaaa g 5361

<210> 20
<211> 1519
<212> DNA
<213> Homo Sapiens

<400> 20
agcaggcgtt tgcgagagga gatacgagct ggacgcctgg cccttccctc ccaccgggtc 60

ctagtccacc gctcccggcg ccggctcccc gcctctcccg ctatgtaccg accgcgagcc 120

cgggcggctc ccgagggcag ggtccggggc tgcgcggtgc ccagcaccgt gctcctgctg 180

ctcgcctacc tggcttacct ggcgctgggc accggcgtgt tctggacgct ggagggccgc 240

gcggcgcagg actccagccg cagcttccag cgcgacaagt gggagctgtt gcagaacttc 300

acgtgtctgg accgcccggc gctggactcg ctgatccggg atgtcgtcca agcatacaaa 360

aacggagcca gcctcctcag caacaccacc agcatggggc gctgggagct cgtgggctcc 420

ttcttctttt ctgtgtccac catcaccacc attggctatg gcaacctgag ccccaacacg 480

atggctgccc gcctcttctg catcttcttt gcccttgtgg ggatcccact caacctcgtg 540

gtgctcaacc gactggggca tctcatgcag cagggagtaa accactgggc cagcaggctg 600

gggggcacct ggcaggatcc tgacaaggcg cggtggctgg cgggctctgg cgccctcctc 660

tcgggcctcc tgctcttcct gctgctgcca ccgctgctct tctcccacat ggagggctgg 720

agctacacag agggcttcta cttcgccttc atcaccctca gcaccgtggg cttcggcgac 780

tacgtgattg gaatgaaccc ctcccagagg tacccactgt ggtacaagaa catggtgtcc 840

ctgtggatcc tctttgggat ggcatggctg gccttgatca tcaaactcat cctctcccag 900

ctggagacgc cagggagggt atgttcctgc tgccaccaca gctctaagga agacttcaag 960

tcccaaagct ggagacaggg acctgaccgg gagccagagt cccactcccc acagcaagga 1020

tgctatccag agggacccat gggaatcata cagcatctgg aaccttctgc tcacgctgca 1080

ggctgtggca aggacagcta gttatactcc attctttggt cgtcgtcctc ggtagcaaga 1140

cccctgattt taagctttgc acatgtccac ccaaactaaa gactacattt tccatccacc 1200

ctagaggctg ggtgcagcta tatgattaat tctgcccaat agggtataca gagacatgtc 1260

ctgggtgaca tgggatgtga ctttcgggtg tcggggcagc atgcccttct cccccacttc 1320

cttactttag cgggctgcaa tgccgccgat atgatggctg ggagctctgg cagccatacg 1380

gcaccatgaa gtagcggcaa tgtttgagcg gcacaattag ataggaagag tctggatctc 1440

tgatgatcac agagccatcc taacaaacgg aatatcaccg accctccttt atgtgagaga 1500

gaaataaaca tctatgaaa 1519

<210> 21
<211> 1832
<212> DNA
<213> Homo Sapiens

<400> 21
aaggacagag gaggggccct tcctgtcagc tggctgggag cagaggtggc tttgtctttt 60

cggaagaact ggttctgtgg aatttgtgct tatttcccat caaggatcaa ggacctgctc 120

tggggctacc tcagggcccc acaggatgag gggctggttt tcagatgagt tttctgcttg 180

cctgtcatct ggatagtgtc taaaaatttg caaactgcct tcttgtcagt gtcttgctca 240

ttcttcatga cactcctgat atgtctctca gtttcctcat ctgctgcctc tccagacttc 300

tgccagaaca ttgcacgcga cagtttcagg cacagaactg actggcagca ggggctgctc 360

cacgagtggg aatttgctcc agcacttcac ggactgcaag cgaggcactt gctaactctt 420

ggataacaag acctctgcca gaagaaccat ggctttggaa ggcggagttc aggctgagga 480

gatgggtgcg gtcctcagtg agcccctgcc tccctgaaca taggaaaccc acctgggcag 540

ccatggaatg ggacaatggc acaggccagg ctctgggctt gccacccacc acctgtgtct 600

accgcgagaa cttcaagcaa ctgctgctgc cacctgtgta ttcggcggtg ctggcggctg 660

gcctgccgct gaacatctgt gtcattaccc agatctgcac gtcccgccgg gccctgaccc 720

gcacggccgt gtacacccta aaccttgctc tggctgacct gctatatgcc tgctccctgc 780

ccctgctcat ctacaactat gcccaaggtg atcactggcc ctttggcgac ttcgcctgcc 840

gcctggtccg cttcctcttc tatgccaacc tgcacggcag catcctcttc ctcacctgca 900

tcagcttcca gcgctacctg ggcatctgcc acccgctggc cccctggcac aaacgtgggg 960

gccgccgggc tgcctggcta gtgtgtgtag ccgtgtggct ggccgtgaca acccagtgcc 1020

tgcccacagc catcttcgct gccacaggca tccagcgtaa ccgcactgtc tgctatgacc 1080

tcagcccgcc tgccctggcc acccactata tgccctatgg catggctctc actgtcatcg 1140

gcttcctgct gcccttgct gccctgctgg cctgctactg tctcctggcc tgccgcctgt 1200

gccgccagga tggcccggca gagcctgtgg cccaggagcg gcgtggcaag gcggcccgca 1260

tggccgtggt ggtggctgct gcctttgcca tcagcttcct gccttttcac atcaccaaga 1320

cagcctacct ggcagtgcgc tcgacgccgg gcgtcccctg cactgtattg gaggcctttg 1380

cagcggccta caaaggcacg cggccgtttg ccagtgccaa cagcgtgctg gaccccatcc 1440

tcttctactt cacccagaag aagttccgcc ggcgaccaca tgagctccta cagaaactca 1500

cagccaaatg gcagaggcag ggtcgctgag tcctccaggt cctgggcagc cttcatattt 1560

gccattgtgt ccggggcacc aggagcccca ccaaccccaa accatgcgga gaattagagt 1620

tcagctcagc tgggcatgga gttaagatcc ctcacaggac ccagaagctc accaaaaact 1680

atttcttcag ccccttctct ggcccagacc ctgtgggcat ggagatggac agacctgggc 1740

ctggctcttg agaggtccca gtcagccatg gagagctggg gaaaccacat taaggtgctc 1800

acaaaaatac agtgtgacgt gtactgtcaa aa 1832

<210> 22
<211> 2811
<212> DNA
<213> Homo Sapiens

<400> 22
acacgtccaa cgccagcatg cagcgcccgg gcccccgcct gtggctggtc ctgcaggtga 60

tgggctcgtg cgccgccatc agctccatgg acatggagcg cccgggcgac ggcaaatgcc 120

agcccatcga gatcccgatg tgcaaggaca tcggctacaa catgactcgt atgcccaacc 180

tgatgggcca cgagaaccag cgcgaggcag ccatccagtt gcacgagttc gcgccgctgg 240

tggagtacgg ctgccacggc cacctccgct tcttcctgtg ctcgctgtac gcgccgatgt 300
gcaccgagca ggtctctacc cccatccccg cctgccgggt catgtgcgag caggcccggc 360
tcaagtgctc cccgattatg gagcagttca acttcaagtg gcccgactcc ctggactgcc 420
ggaaactccc caacaagaac gaccccaact acctgtgcat ggaggcgccc aacaacggct 480
cggacgagcc cacccggggc tcgggcctgt tcccgccgct gttccggccg cagcggcccc 540
acagcgcgca ggagcacccg ctgaaggacg ggggccccgg gcgcggcggc tgcgacaacc 600
cgggcaagtt ccaccacgtg gagaagagcg cgtcgtgcgc gccgctctgc acgcccggcg 660
tggacgtgta ctggagccgc gaggacaagc gcttcgcagt ggtctggctg gccatctggg 720
cggtgctgtg cttcttctcc agcgccttca ccgtgctcac cttcctcatc gacccggccc 780
gcttccgcta ccccgagcgc cccatcatct tcctctccat gtgctactgc gtctactccg 840
tgggctacct catccgcctc ttcgccggcg ccgagagcat cgcctgcgac cgggacagcg 900
gccagctcta tgtcatccag gagggactgg agagcaccgg ctgcacgctg gtcttcctgg 960
tcctctacta cttcggcatg gccagctcgc tgtggtgggt ggtcctcacg ctcacctggt 1020
tcctggccgc cggcaagaag tggggccacg aggccatcga agccaacagc agctacttcc 1080
acctggcagc ctgggccatc ccggcggtga agaccatcct gatcctggtc atgcgcaggg 1140
tggcgggggа cgagctcacc ggggtctgct acgtgggcag catggacgtc aacgcgctca 1200
ccggcttcgt gctcattccc ctggcctgct acctggtcat cggcacgtcc ttcatcctct 1260
cgggcttcgt ggccctgttc cacatccgga gggtgatgaa gacgggcggc gagaacacgg 1320
acaagctgga gaagctcatg gtgcgtatcg ggctcttctc tgtgctgtac accgtgccgg 1380
ccacctgtgt gatcgcctgc tacttttacg aacgcctcaa catggattac tggaagatcc 1440
tggcggcgca gcacaagtgc aaaatgaaca accagactaa aacgctggac tgcctgatgg 1500
ccgcctccat ccccgccgtg gagatcttca tggtgaagat ctttatgctg ctggtggtgg 1560
ggatcaccag cgggatgtgg atttggacct ccaagactct gcagtcctgg cagcaggtgt 1620
gcagccgtag gttaaagaag aagagccgga gaaaaccggc cagcgtgatc accagcggtg 1680
ggatttacaa aaaagcccag catccccaga aaactcacca cgggaaatat gagatccctg 1740
cccagtcgcc cacctgcgtg tgaacagggc tggagggaag ggcacagggg cgcccggagc 1800
taagatgtgg tgcttttctt ggttgtgttt ttctttcttc ttcttctttt ttttttttt 1860
ataaaagcaa aagagaaata cataaaaaag tgtttaccct gaaattcagg atgctgtgat 1920
acactgaaag gaaaaatgta cttaaagggt tttgttttgt tttggttttc cagcgaaggg 1980
aagctcctcc agtgaagtag cctcttgtgt aactaatttg tggtaaagta gttgattcag 2040
ccctcagaag aaaacttttg tttagagccc tccgtaaata tacatctgtg tatttgagtt 2100

ggctttgcta cccatttaca aataagagga cagataactg ctttgcaaat tcaagagcct 2160

cccctgggtt aacaaatgag ccatccccag ggcccacccc caggaaggcc acagtgctgg 2220

gcggcatccc tgcagaggaa agacaggacc cggggcccgc ctcacacccc agtggatttg 2280

gagttgctta aaatagactc tggccttcac caatagtctc tctgcaagac agaaacctcc 2340

atcaaacctc acatttgtga actcaaacga tgtgcaatac attttttttct ctttccttga 2400

aaataaaaag agaaacaagt attttgctat atataaagac aacaaaagaa atctcctaac 2460

aaaagaacta agaggcccag ccctcagaaa cccttcagtg ctacattttg tggcttttta 2520

atggaaacca agccaatgtt atagacgttt ggactgattt gtggaaagga gggggggaaga 2580

gggagaagga tcattcaaaa gttacccaaa gggcttattg actctttcta ttgttaaaca 2640

aatgatttcc acaaacagat caggaagcac taggttggca gagacacttt gtctagtgta 2700

ttctcttcac agtgccagga aagagtggtt tctgcgtgtg tatatttgta atatatgata 2760

tttttcatgc tccactattt tattaaaaat aaaatatgtt ctttaaaaaa a 2811

<210> 23
<211> 2010
<212> DNA
<213> Homo sapiens

<400> 23
ggcagcgact gcgccccgtc ccggcgccgc gctcgtccgc agaggaggcg gcccggcccg 60

ggcagctgcg gctcgggatc cgtcgagggg aggccgagct tgccaagctg gcgcccagcg 120

gggtcatggt gcccggcgcc cgcggcggcg gcgcactggc gcgggctgcc gggcggggcc 180

tcctggcttt gctgctcgcg gtctccgccc cgctccggct gcaggcggag gagctgggtg 240

atggctgtgg acacctagtg acttatcagg atagtggcac aatgacatct aagaattatc 300

ccgggaccta ccccaatcac actgtttgcg aaaagacaat tacagtacca aaggggaaaa 360

gactgattct gaggttggga gatttggata tcgaatccca gacctgtgct tctgactatc 420

ttctcttcac cagctcttca gatcaatatg gtccatactg tggaagtatg actgttccca 480

aagaactctt gttgaacaca agtgaagtaa ccgtccgctt tgagagtgga tcccacattt 540

ctggccgggg ttttttgctg acctatgcga gcagcgacca tccagattta ataacatgtt 600

tggaacgagc tagccattat ttgaagacag aatacagcaa attctgccca gctggttgta 660

gagacgtagc aggagacatt tctgggaata tggtagatgg atatagagat acctctttat 720

tgtgcaaagc tgccatccat gcaggaataa ttgctgatga actaggtggc cagatcagtg 780

tgcttcagcg caaagggatc agtcgatatg aagggattct ggccaatggt gttctttcga 840

gggatggttc cctgtcagac aagcgatttc tgtttacctc caatggttgc agcagatcct 900

tgagttttga acctgacggg caaatcagag cttcttcctc atggcagtcg gtcaatgaga 960

gtggagacca agttcactgg tctcctggcc aagcccgact tcaggaccaa ggcccatcat 1020

gggcttcggg cgacagtagc aacaaccaca aaccacgaga gtggctggag atcgatttgg 1080

gggagaaaaa gaaaataaca ggaattagga ccacaggatc tacacagtcg aacttcaact 1140

tttatgttaa gagttttgtg atgaacttca aaaacaataa ttctaagtgg aagacctata 1200

aaggaattgt gaataatgaa gaaaaggtgt ttcagggtaa ctctaacttt cgggacccag 1260

tgcaaaacaa tttcatccct cccatcgtgg ccagatatgt gcgggttgtc ccccagacat 1320

ggcaccagag gatagccttg aaggtggagc tcattggttg ccagattaca caaggtaatg 1380

attcattggt gtggcgcaag acaagtcaaa gcaccagtgt ttcaactaag aaagaagatg 1440

agacaatcac aaggcccatc ccctcggaag aaacatccac aggaataaac attacaacgg 1500

tggctattcc attggtgctc cttgttgtcc tggtgtttgc tggaatgggg atctttgcag 1560

cctttagaaa gaagaagaag aaaggaagtc cgtatggatc agcagaggct cagaaaacag 1620

actgttggaa gcagattaaa tatccctttg ccagacatca gtcagctgag tttaccatca 1680

gctatgataa tgagaaggag atgacacaaa agttagatct catcacaagt gatatggcag 1740

gttaactccg ttgactgcca aaatagcatc cccaacgtgc agccctccgc atctatcagc 1800

aggttgcccc ggatggatct cagagatgag gatcggaaca ccatgttctt tcccacccta 1860

acaacaacaa agggcagtaa attaaagtac tctttgtaag gtacagttac cgattaatct 1920

agagataaaa tattttctta aaaatatatt tcattaaaca cctatgctgt ctctataaaa 1980

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 2010

<210> 24
<211> 2010
<212> DNA
<213> Homo sapiens

<400> 24
ggcagcgact gcgccccgtc ccggcgccgc gctcgtccgc agaggaggcg gcccggcccg 60

ggcagctgcg gctcgggatc cgtcgagggg aggccgagct tgccaagctg gcgcccagcg 120

gggtcatggt gcccggcgcc cgcggcggcg gcgcactggc gcgggctgcc gggcggggcc 180

tcctggcttt gctgctcgcg gtctccgccc cgctccggct gcaggcggag gagctgggtg 240

atggctgtgg acacctagtg acttatcagg atagtggcac aatgacatct aagaattatc 300

ccgggaccta ccccaatcac actgtttgcg aaaagacaat tacagtacca aaggggaaaa 360

gactgattct gaggttggga gatttggata tcgaatccca gacctgtgct tctgactatc 420

ttctcttcac cagctcttca gatcaatatg gtccatactg tggaagtatg actgttccca 480

aagaactctt gttgaacaca agtgaagtaa ccgtccgctt tgagagtgga tcccacattt 540

ctggccgggg ttttttgctg acctatgcga gcagcgacca tccagattta ataacatgtt 600

tggaacgagc tagccattat ttgaagacag aatacagcaa attctgccca gctggttgta 660

gagacgtagc aggagacatt tctgggaata tggtagatgg atatagagat acctctttat 720

tgtgcaaagc tgccatccat gcaggaataa ttgctgatga actaggtggc cagatcagtg 780

tgcttcagcg caaagggatc agtcgatatg aagggattct ggccaatggt gttctttcga 840

gggatggttc cctgtcagac aagcgatttc tgtttacctc caatggttgc agcagatcct 900

tgagttttga acctgacggg caaatcagag cttcttcctc atggcagtcg gtcaatgaga 960

gtggagacca agttcactgg tctcctggcc aagcccgact tcaggaccaa ggcccatcat 1020

gggcttcggg cgacagtagc aacaaccaca aaccacgaga gtggctggag atcgatttgg 1080

gggagaaaaa gaaaataaca ggaattagga ccacaggatc tacacagtcg aacttcaact 1140

tttatgttaa gagttttgtg atgaacttca aaaacaataa ttctaagtgg aagacctata 1200

aaggaattgt gaataatgaa gaaaaggtgt ttcagggtaa ctctaacttt cgggacccag 1260

tgcaaaacaa tttcatccct cccatcgtgg ccagatatgt gcgggttgtc ccccagacat 1320

ggcaccagag gatagccttg aaggtggagc tcattggttg ccagattaca caaggtaatg 1380

attcattggt gtggcgcaag acaagtcaaa gcaccagtgt ttcaactaag aaagaagatg 1440

agacaatcac aaggcccatc ccctcggaag aaacatccac aggaataaac attacaacgg 1500

tggctattcc attggtgctc cttgttgtcc tggtgtttgc tggaatgggg atctttgcag 1560

cctttagaaa gaagaagaag aaaggaagtc cgtatggatc agcagaggct cagaaaacag 1620

actgttggaa gcagattaaa tatccctttg ccagacatca gtcagctgag tttaccatca 1680

gctatgataa tgagaaggag atgacacaaa agttagatct catcacaagt gatatggcag 1740

gttaactccg ttgactgcca aaatagcatc cccaacgtgc agccctccgc atctatcagc 1800

aggttgcccc ggatggatct cagagatgag gatcggaaca ccatgttctt tcccacccta 1860

acaacaacaa agggcagtaa attaaagtac tctttgtaag gtacagttac cgattaatct 1920

agagataaaa tattttctta aaaatatatt tcattaaaca cctatgctgt ctctataaaa 1980

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 2310

<210> 25
<211> 1159
<212> DNA
<213> Homo Sapiens

<400> 25
agtgccccag gagctatgac aagcaaagga acatacttgc ctggagatag cctttgcgat 60

atttaaatgt ccgtggatac agaaatctct gcaggcaagt tgctccagag catattgcag 120
gacaagcctg taacgaatag ttaaattcac ggcatctgga ttcctaatcc ttttccgaaa 180
tggcaggtgt gagtgcctgt ataaaatatt ctatgtttac cttcaacttc ttgttctggc 240
tatgtggtat cttgatccta gcattagcaa tatgggtacg agtaagcaat gactctcaag 300
caatttttgg ttctgaagat gtaggctcta gctcctacgt tgctgtggac atattgattg 360
ctgtaggtgc catcatcatg attctgggct tcctgggatg ctgcggtgct ataaaagaaa 420
gtcgctgcat gcttctgttg tttttcatag gcttgcttct gatcctgctc ctgcaggtgg 480
cgacaggtat cctaggagct gttttcaaat ctaagtctga tcgcattgtg aatgaaactc 540
tctatgaaaa cacaaagctt ttgagcgcca caggggaaag tgaaaaacaa ttccaggaag 600
ccataattgt gtttcaagaa gagtttaaat gctgcggttt ggtcaatgga gctgctgatt 660
ggggaaataa ttttcaacac tatcctgaat tatgtgcctg tctagataag cagagaccat 720
gccaaagcta taatggaaaa caagtttaca aagagacctg tatttctttc ataaaagact 780
tcttggcaaa aaatttgatt atagttattg gaatatcatt tggactggca gttattgaga 840
tactgggttt ggtgttttct atggtcctgt attgccagat cgggaacaaa tgaatctgtg 900
gatgcatcaa cctatcgtca gtcaaacccc tttaaaatgt tgctttggct ttgtaaattt 960
aaatatgtaa gtgctatata agtcaggagc agctgtcttt ttaaaatgtc tcggctagct 1020
agaccacaga tatcttctag acatattgaa cacatttaag atttgaggga tataagggaa 1080
aatgatatga atgtgtattt ttactcaaaa taaaagtaac tgtttacgtt aaaaaaaaaa 1140
aaaaaaaaaa aaaaaaaaa 1159

<210> 26
<211> 1428
<212> DNA
<213> Homo Sapiens

<400> 26
cttcaggtca gggagaatgt ataaatgtcc attgccatcg aggttctgct atttttgaga 60
agctgaagca actccaagga cacagttcac agaaatttgg ttctcagccc caaaatactg 120
attgaattgg agacaattac aaggactctc tggccaaaaa cccttgaaga ggccccgtga 180
aggaggcagt gaggagcttt tgattgctga cctgtgtcgt accaccccag aatgtgcact 240
gggggctgtg ccagatgcct gggggggacc ctcattcccc ttgctttttt tggcttcctg 300
gctaacatcc tgttattttt tcctggagga aaagtgatag atgacaacga ccacctttcc 360
caagagatct ggtttttcgg aggaatatta ggaagcggtg tcttgatgat cttccctgcg 420
ctggtgttct tgggcctgaa gaacaatgac tgctgtgggt gctgcggcaa cgagggctgt 480

gggaagcgat ttgcgatgtt cacctccacg atatttgctg tggttggatt cttgggagct 540

ggatactcgt ttatcatctc agccatttca atcaacaagg gtcctaaatg cctcatggcc 600

aatagtacat ggggctaccc cttccacgac ggggattatc tcaatgatga ggccttatgg 660

aacaagtgcc gagagcctct caatgtggtt ccctggaatc tgaccctctt ctccatcctg 720

ctggtcgtag gaggaatcca gatggttctc tgcgccatcc aggtggtcaa tggcctcctg 780

gggaccctct gtggggactg ccagtgttgt ggctgctgtg ggggagatgg acccgtttaa 840

acctccgaga tgagctgctc agactctaca gcatgacgac tacaatttct tttcataaaa 900

cttcttctct tcttggaatt attaattcct atctgcttcc tagctgataa agcttagaaa 960

aggcagttat tccttctttc caaccagctt tgctcgagtt agaattttgt tattttcaaa 1020

taaaaaatag tttggccact taacaaattt gatttataaa tctttcaaat tagttccttt 1080

ttagaattta ccaacaggtt caaagcatac ttttcatgat ttttttatta caaatgtaaa 1140

atgtataaag tcacatgtac tgccatacta cttctttgta tataaagatg tttatatctt 1200

tggaagtttt acataaatca aaggaagaaa gcacatttaa aatgagaaac taagaccaat 1260

ttctgttttt aagaggaaaa agaatgattg atgtatccta agtattgtta tttgttgtct 1320

ttttttgctg ccttgcttga gttgcttgtg actgatcttt tgaggctgtc atcatggcta 1380

gggttctttt atgtatgtta aattaaaacc tgaattcaga ggtaacgt 1428

<210> 27
<211> 2454
<212> DNA
<213> Homo Sapiens

<400> 27
ccaggaggca cgctggtttt ccggggccgc tccatcgcgc cttcctcctg cgcctcgctt 60

ctccggtcca gccgccatct tcctttccgc acaggggccg ccgagcgggg ccatgcagcc 120

aacgctgctt ctcagcctcc tgggagccgt ggggctggcg gctgtcaatt ccatgccagt 180

ggataacagg aaccacaatg aaggaatggt gactcgctgc atcattgagg tcctctcaaa 240

tgccttgtcg aagtccagcg ctccacccat caccctgag tgccgccaag tcctgaagac 300

gagtagaaaa gacgtcaaag acaaagagac aactgaaaat gaaaacacaa agtttgaagt 360

aagattgtta agagacccag ctgatgcctc ggaagcccac gagtcctcca gcaggggaga 420

ggcaggagcc ccaggggagg aggacatcca aggcccaaca aaggcagaca cagagaaatg 480

ggcagaggga ggcgggcaca gccgagagcg agcggatgag ccccagtgga gcctctatcc 540

ctccgacagc caagtctctg aagaagtgaa gacacgccat tctgagaaga gccagagaga 600

ggatgaggag gaggaggagg gagagaacta tcaaaaaggg gagcgagggg aagatagcag 660

tgaagagaaa caccttgaag agccaggaga gacacaaaac gcttttctca atgaaagaaa 720
gcaggcttca gctataaaaa aagaggagtt agtggccaga tcggaaacac atgctgccgg 780
gcattctcag gagaagacac atagccgaga gaagagtagc caggagagtg gagaggaggc 840
agggagccag gagaatcacc cccaggagtc taaaggccaa ccccgaagcc aggaagaatc 900
tgaggaaggt gaggaagatg ccacctctga ggtggacaaa cgacgcacga ggcccagaca 960
ccaccacggg aggagcaggc ccgacaggtc ctctcaagga gggagtcttc cctctgagga 1020
aaagggacac ccccaggagg aatctgagga gtcaaacgtc agcatggcca gtttagggga 1080
aaagagggac caccattcaa cccactacag ggcttcagag gaagaacctg aatatggaga 1140
agaaataaag ggttatccag gcgtccaggc ccctgaggac ctggagtggg agcgctatag 1200
gggcagagga agtgaagaat acagggctcc aagacctcag agtgaggaga gttgggatga 1260
ggaggacaag agaaactacc ccagcttaga gcttgataag atggcacatg gatatggtga 1320
agaaagtgag gaagagaggg gccttgagcc gggaaaggga cgccatcaca gaggcagggg 1380
aggggagcca cgtgcctatt tcatgtctga caccagagaa gagaaaaggt tcttgggtga 1440
aggacaccac cgtgtccaag aaaaccagat ggacaaggca aggaggcatc cacaaggtgc 1500
gtggaaagag ctggacagaa attatctcaa ctacggtgag gaaggagccc cagggaagtg 1560
gcagcagcag ggagacctgc aggacactaa agaaaacagg gaggaagcta ggtttcaaga 1620
taaacaatat agctcccatc acacagctga aaagaggaag agattagggg aactgttcaa 1680
cccatactac gaccctctcc agtggaagag cagccatttt gaaagaagag acaacatgaa 1740
tgacaatttt ctcgagggtg aggaggaaaa tgagctgacc ttgaacgaga agaatttctt 1800
cccagaatac aactatgact ggtgggagaa aaagcccttc tctgaggatg tgaactgggg 1860
gtatgagaag agaaacctcg ccagggtccc caagctggac ctgaaaaggc aatatgacag 1920
ggtggcccaa ctggaccagc tccttcacta caggaagaag tcagctgagt ttccagactt 1980
ctatgattct gaggagccgg tgagcaccca ccaggaggca gaaaatgaaa aggacagggc 2040
tgaccagaca gtcctgacag aggacgagaa aaaagaactc gaaaacttgg ctgcaatgga 2100
tttggaacta cagaagatag ctgagaaatt cagccaaagg ggctgactgt cattggagcg 2160
gtgggcactg ttaagaagca gccatcacat gatctgtttt tcaccacttc actgaaagac 2220
accatttata tacccaaggg cagaaagtag aacttactat tcattaaatg tttgacacaa 2280
ttggaattgt ctttaatttc tgtcagaatg ctattgaaaa tgtgaattgc atgacttgta 2340
gcatattctt ttctgcaaaa tagacatatt aacatgctta tgacaatgac tgtgctactg 2400
tctttggaaa aatgtttgtc tcagttggaa ataataaaag attcacctga gacc 2454

<210> 28
<211> 1980
<212> DNA
<213> Homo Sapiens

<400> 28
cttcttgtgg tagggacctc tcctcagtat ttgaaactaa ccagcatctg acagatttcg 60

aatttgtaaa aaataccctc gaagattcag gaatgaagct tctgtgtgaa ggattaaaac 120

agcccaactg tgtattacag acattgaggt ggtaccggtg ccttatctct tctgcttctt 180

gtggggctct agcagctgtt cttagcacca gtcagtggct cactgaactg gaatttagtg 240

agacaaaact ggaagcttca gctttgaaat tgctctatgg aggcttaaaa gatccaaatt 300

gcaaattaca gaagctcaac ttgcagtttt ctttatctgt aaccgctgca aaacttccag 360

ttggaatggt tggaaattgt tctggtttct cgggatcatt ggtgcaatct cattttggct 420

actgtcagga cagttctttc aaatgtgatc tttgtaagct gctctggcct tccaccagag 480

ttgctgctgc aaaggattgt gggagtccta agtccttcct atcagaaggg ctgaactggg 540

caggaagact tgaggcagtg gaggaggttt tggggttggg ggtgcttgta cagcccggtg 600

acccagcatc tcagggtggg gggcattgtg aaaactatgg gtcttttaga gacttggtgg 660

acttagaagt caaggcagaa ccaagcctga gaaaaggtgg tatggatctc cagagaccca 720

ccctacaagt tgtcctcctt tgcaaaatct tctccctcaa actatttctc tttattgcat 780

tgcctaattc tcctggtcag gttagtgtgg tgcaagtgac catcccagac ggtttcgtga 840

acgtgactgt tggatctaat gtcactctca tctgcatcta caccaccact gtggcctccc 900

gagaacagct ttccatccag tggtctttct tccataagaa ggagatggag ccaatttctt 960

ctccttggga ggaggggaag tggccagatg ttgaggctgt gaagggcact cttgatggac 1020

agcaggctga actccagatt tacttttctc aaggtggaca agctgtagcc atcgggcaat 1080

ttaaagatcg aattacaggg tccaacgatc caggtaatgc atctatcact atctcgcata 1140

tgcagccagc agacagtgga atttacatct gcgatgttaa caacccccca gactttctcg 1200

gccaaaacca aggcatcctc aacgtcagtg tgttagtgaa accttctaag ccccttttgta 1260

gcgttcaagg aagaccagaa actggccaca ctatttccct ttcctgtctc tctgcgcttg 1320

gaacaccttc ccctgtgtac tactggcata aacttgaggg aagagacatc gtgccagtga 1380

aagaaaactt caacccaacc accgggattt tggtcattgg aaatctgaca aattttgaac 1440

aaggttatta ccagtgtact gccatcaaca gacttggcaa tagttcctgc gaaatcgatc 1500

tcacttcttc acatccagaa gttggaatca ttgttggggc cttgattggt agcctggtag 1560

gtgccgccat catcatctct gttgtgtgct tcgcaaggaa taaggcaaaa gcaaaggcaa 1620

aagaaagaaa ttctaagacc atcgcggaac ttgagccaat gacaaagata aacccaaggg 1680

gagaaagcga agcaatgcca agagaagacg ctacccaact agaagtaact ctaccatctt 1740

ccattcatga gactggccct gataccatcc aagaaccaga ctatgagcca aagcctactc 1800

aggagcctgc cccagagcct gccccaggat cagagcctat ggcagtgcct gaccttgaca 1860

tcgagctgga gctggagcca gaaacgcagt cggaattgga gccagagcca gagccagagc 1920

cagagtcaga gcctggggtt gtagttgagc ccttaagtga agatgaaaag ggagtggtta 1980

<210> 29
<211> 1242
<212> DNA
<213> Homo Sapiens

<400> 29
atggtgttcg cattttggaa ggtctttctg atcctaagct gccttgcagg tcaggttagt 60

gtggtgcaag tgaccatccc agacggtttc gtgaacgtga ctgttggatc taatgtcact 120

ctcatctgca tctacaccac cactgtggcc tcccgagaac agctttccat ccagtggtct 180

ttcttccata agaaggagat ggagccaatt tcttctcctt gggaggaggg gaagtggcca 240

gatgttgagg ctgtgaaggg cactcttgat ggacagcagg ctgaactcca gatttacttt 300

tctcaaggtg gacaagctgt agccatcggg caatttaaag atcgaattac agggtccaac 360

gatccaggta atgcatctat cactatctcg catatgcagc cagcagacag tggaatttac 420

atctgcgatg ttaacaaccc cccagacttt ctcggccaaa accaaggcat cctcaacgtc 480

agtgtgttag tgaaaccttc taagcccctt tgtagcgttc aaggaagacc agaaactggc 540

cacactattt ccctttcctg tctctctgcg cttggaacac cttcccctgt gtactactgg 600

cataaacttg agggaagaga catcgtgcca gtgaaagaaa acttcaaccc aaccaccggg 660

attttggtca ttggaaatct gacaaatttt gaacaaggtt attaccagtg tactgccatc 720

aacagacttg gcaatagttc ctgcgaaatc gatctcactt cttcacatcc agaagttgga 780

atcattgttg gggccttgat tggtagcctg gtaggtgccg ccatcatcat ctctgttgtg 840

tgcttcgcaa ggaataaggc aaaagcaaag gcaaaagaaa gaaattctaa gaccatcgcg 900

gaacttgagc caatgacaaa gataaaccca aggggagaaa gcgaagcaat gccaagagaa 960

gacgctaccc aactagaagt aactctacca tcttccattc atgagactgg ccctgatacc 1020

atccaagaac cagactatga gccaaagcct actcaggagc ctgccccaga gcctgcccca 1080

ggatcagagc ctatggcagt gcctgacctt gacatcgagc tggagctgga gccagaaacg 1140

cagtcggaat tggagccaga gccagagcca gagccagagt cagagcctgg ggttgtagtt 1200

gagcccttaa gtgaagatga aaagggagtg gttaaggcat ag 1242

<210> 30
<211> 1451
<212> DNA
<213> Homo Sapiens

<400> 30
cggcccgccc tggggaggcg cgcagcagag gctccgattc ggggcaggtg agaggctgac 60

tttctctcgg tgcgtccagt ggagctctga gtttcgaatc ggtggcggcg gattccccgc 120

gcgcccggcg tcggggcttc caggaggatg cggagcccca gcgcggcgtg gctgctgggg 180

gccgccatcc tgctagcagc ctctctctcc tgcagtggca ccatccaagg aaccaataga 240

tcctctaaag gaagaagcct tattggtaag gttgatggca catcccacgt cactggaaaa 300

ggagttacag ttgaaacagt cttttctgtg gatgagtttt ctgcatctgt cctcactgga 360

aaactgacca cggtcttcct tccaattgtc tacacaattg tgtttgtggt gggtttgcca 420

agtaacggca tggccctgtg ggtctttctt ttccgaacta agaagaagca ccctgctgtg 480

atttacatgg ccaatctggc cttggctgac ctcctctctg tcatctggtt ccccttgaag 540

attgcctatc acatacatgc caacaactgg atttatgggg aagctctttg taatgtgctt 600

attggctttt tctatggcaa catgtactgt tccattctct tcatgacctg cctcagtgtg 660

cagaggtatt gggtcatcgt gaaccccatg gggcactcca ggaagaaggc aaacattgcc 720

attggcatct ccctggcaat atggctgctg attctgctgg tcaccatccc tttgtatgtc 780

gtgaagcaga ccatcttcat tcctgccctg aacatcacga cctgtcatga tgttttgcct 840

gagcagctct tggtgggaga catgttcaat tacttcctct ctctggccat tggggtcttt 900

ctgttcccag ccttcctcac agcctctgcc tatgtgctga tgatcagaat gctgcgatct 960

tctgccatgg atgaaaactc agagaagaaa aggaagaggg ccatcaaact cattgtcact 1020

gtcctggcca tgtacctgat ctgcttcact cctagtaacc ttctgcttgt ggtgcattat 1080

tttctgatta agagccaggg ccagagccat gtctatgccc tgtacattgt agccctctgc 1140

ctctctaccc ttaacagctg catcgacccc tttgtctatt actttgtttc acatgatttc 1200

agggatcatg caaagaacgc tctcctttgc cgaagtgtcc gcactgtaaa gcagatgcaa 1260

gtatccctca cctcaaagaa acactccagg aaatccagct cttactcttc aagttcaacc 1320

actgttaaga cctcctattg agttttccag gtcctcagat gggaattgca cagtaggatg 1380

tggaacctgt ttaatgttat gaggacgtgt ctgttatttc ctaatcaaaa aggtctcacc 1440

acataccacc g 1451

<210> 31
<211> 5115
<212> DNA

<213> Homo Sapiens

<400> 31
gaattccggg agcgggcggg ctgcgaggcc gcggggcatg cgggaggcgg aggggtggga 60
ccgggtggct gcgcccattc cacacccgcc gaaagcggac actgtcagct gaatcactcc 120
ccttttagga ggagggaggg ggaaaaggtg tctagctaat ttctgcttaa aaaagcacag 180
gagatcgcgg gtcagctttg cagtcgctgc cttctcgcgc ctgaccatgc acccctgcat 240
cttcctgctg ggcacaggcg agcgctttat ttctggagct gagggctaaa actttttca 300
cttttcttct cctcaacatc tgaatcatgc catgtgccca gaggagctgg cttgcaaacc 360
tttccgtggt ggctcagctc cttaactttg gggcgctttg ctatgggaga cagcctcagc 420
caggcccggt tcgcttcccg gacaggaggc aagagcattt tatcaagggc ctgccagaat 480
accacgtggt gggtccagtc cgagtagatg ccagtgggca ttttttgtca tatggcttgc 540
actatcccat cacgagcagc aggaggaaga gagatttgga tggctcagag gactgggtgt 600
actacagaat ttctcacgag gagaaggacc tgtttttttaa cttgacggtc aatcaaggat 660
ttctttccaa tagctacatc atggagaaga gatatgggaa cctctcccat gttaagatga 720
tggcttcctc tgccccctc tgccatctca gtggcacggt tctacagcag ggcaccagag 780
ttgggacggc agccctcagt gcctgccatg gactgactgg atttttccaa ctaccacatg 840
gagacttttt cattgaaccc gtgaagaagc atccactggt tgagggaggg taccacccgc 900
acatcgttta caggaggcag aaagttccag aaaccaagga gccaacctgt ggattaaagg 960
acagtgttaa catctcccag aagcaagagc tatggcggga gaagtgggag aggcacaact 1020
tgccaagcag aagcctctct cggcgttcca tcagcaagga gagatgggtg gagacactgg 1080
tggtggccga cacaaagatg attgaatacc atgggagtga gaatgtggag tcctacatcc 1140
tcaccatcat gaacatggtc actgggttgt tccataaccc aagcattggc aatgcaattc 1200
acattgttgt ggttcggctc attctactcg aagaagaaga gcaaggactg aaaatagttc 1260
accatgcaga aaagacactg tctagcttct gcaagtggca gaagagtatc aatcccaaga 1320
gtgacctcaa tcctgttcat cacgacgtgg ctgtccttct caccagaaag gacatctgtg 1380
ctggtttcaa tcgcccctgc gagaccctgg gcctgtctca cctttcagga atgtgtcagc 1440
ctcaccgcag ttgtaacatc aatgaagatt cgggactccc tctggctttc acaattgccc 1500
atgagctagg acacagcttc ggcatccagc atgatgggaa agaaaatgac tgtgagcctg 1560
tgggcagaca tccgtacatc atgtcccgcc agctccagta cgatcccact ccgctgacat 1620
ggtccaagtg cagcgaggag tacatcaccc gcttcttgga ccgaggctgg gggttctgtc 1680
ttgatgacat acctaaaaag aaaggcttga agtccaaggt cattgccccc ggagtgatct 1740

atgatgttca ccaccagtgc cagctacaat atggacccaa tgctaccttc tgccaggaag 1800

tagaaaacgt ctgccagaca ctgtggtgct ccgtgaaggg cttttgtcgc tctaagctgg 1860

acgctgctgc agatggaact caatgtggtg agaagaagtg gtgtatggca ggcaagtgca 1920

tcacagtggg gaagaaacca gagagcattc ctggaggctg gggccgctgg tcaccctggt 1980

cccactgttc caggacctgt ggggctggag tccagagcgc agagaggctc tgcaacaacc 2040

ccgagccaaa gtttggaggg aaatattgca ctggagaaag aaaacgctat cgcttgtgca 2100

acgtccaccc ctgtcgctca gaggcaccaa catttcggca gatgcagtgc agtgaatttg 2160

acactgttcc ctacaagaat gaactctacc actggtttcc catttttaac ccagcacatc 2220

cttgtgagct ctactgccga cccatagatg gccagttttc tgagaaaatg ctggatgctg 2280

tcattgatgg tacccccttgc tttgaaggcg gcaacagcag aaatgtctgt attaatggca 2340

tatgtaagat ggttggctgt gactatgaga tcgattccaa tgccaccgag gatcgctgcg 2400

gtgtgtgcct gggagatggc tcttcctgcc agactgtgag aaagatgttt aagcagaagg 2460

aaggatctgg ttatgttgac attgggctca ttccaaaagg agcaagggac ataagagtga 2520

tggaaattga gggagctgga aacttcctgg ccatcaggag tgaagatcct gaaaaatatt 2580

acctgaatgg agggtttatt atccagtgga acgggaacta taagctggca gggactgtct 2640

ttcagtatga caggaaagga gacctggaaa agctgatggc cacaggtccc accaatgagt 2700

ctgtgtggat ccagcttcta ttccaggtga ctaaccctgg catcaagtat gagtacacaa 2760

tccagaaaga tggccttgac aatgatgttg agcagatgta cttctggcag tacggccact 2820

ggacagagtg cagtgtgacc tgcgggacag gtatccgccg ccaaactgcc cattgcataa 2880

agaagggccg cgggatggtg aaagctacat tctgtgaccc agaaacacag cccaatggga 2940

gacagaagaa gtgccatgaa aaggcttgtc cacccaggtg gtgggcaggg gagtgggaag 3000

catgctcggc gacatgcggg ccccacgggg agaagaagcg aaccgtgctg tgcatccaga 3060

ccatggtctc tgacgagcag gctctcccgc ccacagactg ccagcacctg ctgaagccca 3120

agaccctcct ttcctgcaac agagacatcc tgtgcccctc ggactggaca gtgggcaact 3180

ggagtgagtg ttctgtttcc tgtggtggtg gagtgcggat tcgcagtgtc acatgtgcca 3240

agaaccatga tgaaccttgc gatgtgacaa ggaaacccaa cagccgagct ctgtgtggcc 3300

tccagcaatg cccttctagc cggagagttc tgaaaccaaa caaaggcact atttccaatg 3360

gaaaaaaccc accaacacta aagcccgtcc ctccacctac atccaggccc agaatgctga 3420

ccacacccac agggcctgag tctatgagca caagcactcc agcaatcagc agccctagtc 3480

ctaccacagc ctccaaagaa ggagacctgg gtgggaaaca gtggcaagat agctcaaccc 3540

aacctgagct gagctctcgc tatctcattt ccactggaag cacttcccag cccatcctca 3600

cttcccaatc cttgagcatt cagccaagtg aggaaaatgt ttccagttca gatactggtc 3660

ctacctcgga gggaggcctt gtagctacaa caacaagtgg ttctggcttg tcatcttccc 3720

gcaaccctat cacttggcct gtgactccat tttacaatac cttgaccaaa ggtccagaaa 3780

tggagattca cagtggctca ggggaagaaa gagaacagcc tgaggacaaa gatgaaagca 3840

atcctgtaat atggaccaag atcagagtac ctggaaatga cgctccagtg gaaagtacag 3900

aaatgccact tgcacctcca ctaacaccag atctcagcag ggagtcctgg tggccaccct 3960

tcagcacagt aatggaagga ctgctcccca gccaaaggcc cactacttcc gaaactggga 4020

cacccagagt tgaggggatg gttactgaaa agccagccaa cactctgctc cctctgggag 4080

gagaccacca gccagaaccc tcaggaaaga cggcaaaccg taaccacctg aaacttccaa 4140

acaacatgaa ccaaacaaaa agttctgaac cagtcctgac tgaggaggat gcaacaagtc 4200

tgattactga gggctttttg ctaaatgcct ccaattacaa gcagctcaca aacggccacg 4260

gctctgcaca ctggatcgtc ggaaactgga gcgagtgctc caccacatgt ggcctggggg 4320

cctactggaa aagggtggag tgcaccaccc agatggattc tgactgtgcg gccatccaga 4380

gacctgaccc tgcaaaaaga tgccacctcc gtccctgtgc tggctggaaa gtgggaaact 4440

ggagcaagtg ctccagaaac tgcagtgggg gcttcaagat acgcgagatt cagtgcgtgg 4500

acagccggga ccaccggaac ctgaggccat ttcactgcca gttcctggcc ggcattcctc 4560

ccccattgag catgagctgt aacccggagc cctgtgaggc gtggcaggtg gagccttgga 4620

gccagtgctc caggtcctgt ggaggtggag ttcaggagag aggagtgttc tgtccaggag 4680

gcctctgtga ttggacaaaa agacccacat ccaccatgtc ttgcaatgag cacctgtgct 4740

gtcactgggc cactgggaac tgggacctgt gttccacttc ctgtggaggt ggctttcaga 4800

agaggattgt ccaatgtgtg ccctcagagg gcaataaaac tgaagaccaa gaccaatgtc 4860

tatgtgatca caaacccaga cctccagaat tcaaaaaatg caaccagcag gcctgcaaga 4920

aaagtgccga tttactttgc actaaggaca aactgtcagc cagtttctgc cagacactga 4980

aagccatgaa gaaatgttct gtgcccaccg tgagggctga gtgctgcttc tcgtgtcccc 5040

agacacacat cacacacacc caaaggcaaa gaaggcaacg gttgctccaa aagtcaaaag 5100

aactctaagc ccaaa 5115

<210> 32
<211> 799
<212> DNA
<213> Homo Sapiens

<400> 32
cactcccaaa gaactgggta ctcaacactg agcagatctg ttctttgagc taaaaaccat 60

gtgctgtacc aagagtttgc tcctggctgc tttgatgtca gtgctgctac tccacctctg 120

cggcgaatca gaagcagcaa gcaactttga ctgctgtctt ggatacacag accgtattct 180

tcatcctaaa tttattgtgg gcttcacacg gcagctggcc aatgaaggct gtgacatcaa 240

tgctatcatc tttcacacaa agaaaaagtt gtctgtgtgc gcaaatccaa aacagacttg 300

ggtgaaatat attgtgcgtc tcctcagtaa aaaagtcaag aacatgtaaa aactgtggct 360

tttctggaat ggaattggac atagcccaag aacagaaaga accttgctgg ggttggaggt 420

ttcacttgca catcatggag ggtttagtgc ttatctaatt tgtgcctcac tggacttgtc 480

caattaatga agttgattca tattgcatca tagtttgctt tgtttaagca tcacattaaa 540

gttaaactgt attttatgtt atttatagct gtaggttttc tgtgtttagc tatttaatac 600

taattttcca taagctattt tggtttagtg caaagtataa aattatattt gggggggaat 660

aagattatat ggactttctt gcaagcaaca agctattttt taaaaaaact atttaacatt 720

cttttgttta tattgttttg tctcctaaat tgttgtaatt gcattataaa ataagaaaaa 780

cattaataag acaaatatt 799

<210> 33
<211> 1901
<212> DNA
<213> Homo Sapiens

<400> 33
gggcaggaag acggcgctgc ccggaggagc ggggcgggcg ggcgcgcggg ggagcgggcg 60

gcgggcggga gccaggcccg ggcgggggcg ggggcggcgg ggccagaaga ggcggcgggc 120

cgcgctccgg ccggtctgcg gcgttggcct tggctttggc tttggcggcg gcggtggaga 180

agatgctgca gtccctggcc ggcagctcgt gcgtgcgcct ggtggagcgg caccgctcgg 240

cctggtgctt cggcttcctg gtgctgggct acttgctcta cctggtcttc ggcgcagtgg 300

tcttctcctc ggtggagctg ccctatgagg acctgctgcg ccaggagctg cgcaagctga 360

agcgacgctt cttggaggag cacgagtgcc tgtctgagca gcagctggag cagttcctgg 420

gccgggtgct ggaggccagc aactacggcg tgtcggtgct cagcaacgcc tcgggcaact 480

ggaactggga cttcacctcc gcgctcttct tcgccagcac cgtgctctcc accacaggtt 540

atggccacac cgtgcccttg tcagatggag gtaaggcctt ctgcatcatc tactccgtca 600

ttggcattcc cttcaccctc ctgttcctga cggctgtggt ccagcgcatc accgtgcacg 660

tcacccgcag gccggtcctc tacttccaca tccgctgggg cttctccaag caggtggtgg 720

ccatcgtcca tgccgtgctc cttgggtttg tcactgtgtc ctgcttcttc ttcatcccgg 780

ccgctgtctt ctcagtcctg gaggatgact ggaacttcct ggaatccttt tatttttgtt 840

ttatttccct gagcaccatt ggcctggggg attatgtgcc tggggaaggc tacaatcaaa 900

aattcagaga gctctataag attgggatca cgtgttacct gctacttggc cttattgcca 960

tgttggtagt tctggaaacc ttctgtgaac tccatgagct gaaaaaattc agaaaaatgt 1020

tctatgtgaa gaaggacaag gacgaggatc aggtgcacat catagagcat gaccaactgt 1080

ccttctcctc gatcacagac caggcagctg gcatgaaaga ggaccagaag caaaatgagc 1140

cttttgtggc cacccagtca tctgcctgcg tggatggccc tgcaaaccat tgagcgtagg 1200

atttgttgca ttatgctaga gcaccagggt cagggtgcaa ggaagaggct taagtatgtt 1260

catttttatc agaatgcaaa agcgaaaatt atgtcacttt aagaaatagc tactgtttgc 1320

aatgtcttat taaaaaacaa caaaaaaaga cacatggaac aaagaagctg tgaccccagc 1380

aggatgtcta atatgtgagg aaatgagatg tccacctaaa attcatatgt gacaaaatta 1440

tctcgacctt acataggagg agaatacttg aagcagtatg ctgctgtggt tagaagcaga 1500

ttttatactt ttaactggaa actttggggt ttgcatttag atcatttagc tgatggctaa 1560

atagcaaaat ttatatttag aagcaaaaaa aaaaagcata gagatgtgtt ttataaatag 1620

gtttatgtgt actggtttgc atgtacccac ccaaaatgat tatttttgga gaatctaagt 1680

caaactcact atttataatg cataggtaac cattaactat gtacatataa agtataaata 1740

tgtttatatt ctgtacatat ggtttaggtc accagatcct agtgtagttc tgaaactaag 1800

actatagata ttttgtttct tttgatttct ctttatacta aagaatccag agttgctaca 1860

ataaaataag gggaataata aacttgagag tgaataacca t 1901

<210> 34
<211> 492
<212> DNA
<213> Homo Sapiens

<400> 34
aaagggactc cttgaaactg attgagagcc cagtggattt gccagcagtt tgagcttcta 60

ccgagtcttc ccccacctca atccctgttg ctatggagac taccaatgga acggagacct 120

ggtatgagag cctgcatgcc gtgctgaagg ctctaaatgc cactcttcac agcaatttgc 180

tctgccggcc agggccaggg ctggggccag acaaccagac tgaagagagg cgggccagcc 240

tacctggccg tgatgacaac tcctacatgt acattctctt tgtcatgttt ctatttgctg 300

taactgtggg cagcctcatc ctgggataca cccgctcccg caaagtggac aagcgtagtg 360

acccctatca tgtgtatatc aagaaccgtg tgtctatgat ctaacacgag agggctggga 420

cggtggaaga ccaagacacc tggggattgc gtctggggcc tccagaactc tgctgtggac 480

tgcatcaggt ct 492

<210> 35
<211> 14756
<212> DNA
<213> Homo Sapiens

<400> 35
ctgggcggcc gggcgcgggg agagggcgcg ggagcggctc gtgcggcagg taccatgcgg 60

acgcgcgagc ccggcgaggc cccggcaggc ccgtccctgc tcgggggcgc gctgagacgg 120

cgggtgagct ccacgagagc gccgtcgcca cttcgggcca actttgcgat tcccgacagt 180

taagcaatgg ggagacattt ggctttgctc ctgcttctgc tccttctctt ccaacatttt 240

ggagacagtg atggcagcca acgacttgaa cagactcctc tgcagtttac acacctcgag 300

tacaacgtca ccgtgcagga gaactctgca gctaagactt atgtggggca tcctgtcaag 360

atgggtgttt acattacaca tccagcgtgg gaagtaaggt acaaaattgt ttccggagac 420

agtgaaaacc tgttcaaagc tgaagagtac attctcggag acttttgctt tctaagaata 480

aggaccaaag gaggaaatac agctattctt aatagagaag tgaaggatca ctacacattg 540

atagtgaaag cacttgaaaa aaatactaat gtggaggcgc gaacaaaggt cagggtgcag 600

gtgctggata caaatgactt gagaccgtta ttctcaccca cctcatacag cgtttcttta 660

cctgaaaaca cagctataag gaccagtatc gcaagagtca gcgccacgga tgcagacata 720

ggaaccaacg gggaatttta ctacagtttt aaagatcgaa cagatatgtt tgctattcac 780

ccaaccagtg gtgtgatagt gttaactggt agacttgatt acctagagac caagctctat 840

gagatggaaa tcctcgctgc ggaccgtggc atgaagttgt atgggagcag tggcatcagc 900

agcatggcca agctaacggt gcacatcgaa caggccaatg aatgtgctcc ggtgataaca 960

gcagtgacat tgtcaccatc agaactggac agggacccag catatgcaat tgtgacagtg 1020

gatgactgcg atcagggtgc caatggtgac atagcatctt taagcatcgt ggcaggtgac 1080

cttctccagc agtttagaac agtgaggtcc tttccaggga gtaaggagta taaagtcaaa 1140

gccatcggtg acattgattg ggacagtcat cctttcggct acaatctcac actacaggct 1200

aaagataaag gaactccgcc ccagttctct tctgttaaag tcattcacgt gacttctcca 1260

cagttcaaag ccgggccagt caagtttgaa aaggatgttt acagagcaga aataagtgaa 1320

tttgctcctc ccaacacacc tgtggtcatg gtaaaggcca ttcctgctta ttcccatttg 1380

aggtatgttt ttaaaaggac acctggaaaa gctaaattca gtttaaatta caacactggt 1440

ctcatttcta ttttagaacc agttaaaaga cagcaggcag cccattttga acttgaagta 1500

acaacaagtg acagaaaagc gtccaccaag gtcttggtga aagtcttagg tgcaaatagc 1560

aatccccctg aatttaccca gacagcgtac aaagctgctt ttgatgagaa cgtgcccatt 1620

ggtactacta tcatgagcct gagtgccgta gaccctgatg agggtgagaa tgggtacgtg 1680

acatacagta tcgcaaattt aaatcatgtg ccgtttgcga ttgaccattt cactggtgcc 1740

gtgagtacgt cagaaaacct ggactacgaa ctgatgcctc gggtttatac tctgaggatt 1800

cgtgcatcag actggggctt gccgtaccgc cgggaagtcg aagtccttgc tacaattact 1860

ctcaataact tgaatgacaa cacacctttg tttgagaaaa taaattgtga agggacaatt 1920

cccagagatc taggcgtggg agagcaaata accactgttt ctgctattga tgcagatgaa 1980

cttcagttgg tacagtatca gattgaagct ggaaatgaac tggatttgtt tagtttaaac 2040

cccaactcgg gggtattgtc attaaagcga tcgctaatgg atggcttagg tgcaaaggtg 2100

tctttccaca gtctgagaat cacagctaca gatggagaaa attttgccac accattatat 2160

atcaacataa cagtggctgc cagtcacaag ctggtaaact tgcagtgtga agagactggt 2220

gttgccaaaa tgctggcaga gaagctcctg caggcaaata aattacacaa ccagggagag 2280

gtggaggata ttttcttcga ttctcactct gtcaatgctc acataccgca gtttagaagc 2340

actcttccga ctggtattca ggtaaaggaa aaccagcctg tgggttccag tgtaattttc 2400

atgaactcca ctgaccttga cactggcttc aatggaaaac tggtctatgc tgtttctgga 2460

ggaaatgagg atagttgctt catgattgat atggaaacag gaatgctgaa aattttatct 2520

cctcttgacc gtgaaacaac agacaaatac accctgaata ttaccgtcta tgaccttggg 2580

ataccccaga aggctgcgtg gcgtcttcta catgtcgtgg ttgtcgatgc caatgataat 2640

ccacccgagt ttttacagga gagctatttt gtggaagtga gtgaagacaa ggaggtacat 2700

agtgaaatca tccaggttga agccacagat aaagacctgg ggcccaacgg acacgtgacg 2760

tactcaattc ttacagacac agacacattt tcaattgaca gcgtgacggg tgttgttaac 2820

atcgcacgcc ctctggatcg agagctgcag catgagcact ccttaaagat tgaggccagg 2880

gaccaagcca gagaagagcc tcagctgttc tccactgtcg ttgtgaaagt atcactagaa 2940

gatgttaatg acaacccacc tacatttatt ccacctaatt atcgtgtgaa agtccgagag 3000

gatcttccag aaggaaccgt catcatgtgg ttagaagccc acgatcctga tttaggtcag 3060

tctggtcagg tgagatacag ccttctggac cacggagaag gaaacttcga tgtggataaa 3120

ctcagtggag cagttaggat cgtccagcag ttggactttg agaagaagca agtgtataat 3180

ctcactgtga gggccaaaga caagggaaag ccagtttctc tgtcttctac ttgctatgtt 3240

gaagttgagg tggttgatgt gaatgagaac ctgcacccac ccgtgttttc cagctttgtg 3300

gaaaagggga cagtgaaaga agatgcacct gttggttcat tggtaatgac ggtgtcggct 3360

catgatgagg acgccggaag agatggggag atccgatact ccattagaga tggctctggc 3420

gttggtgttt tcaaaatagg tgaagagaca ggtgtcatag agacgtcaga tcgactggac 3480
cgtgaatcga cctcccatta ttggctaaca gtctttgcaa ccgatcaggg tgtcgtgcct 3540
ctttcatcgt tcatagagat ctacatagag gttgaggatg tcaatgacaa tgcaccacag 3600
acatcagagc ctgtttatta cccagaaatc atggaaaatt ctcctaaaga tgtatctgtg 3660
gtccagatcg aggcatttga tccagattcg agctctaatg acaagctcat gtacaaaatt 3720
acaagtggaa atccacaagg attcttttca atacatccta aaacaggtct catcacaact 3780
acgtcaagga agctagaccg agaacagcaa gatgaacaca tattagaggt tactgtgaca 3840
gacaatggta gtccccccaa atcaaccatt gcaagagtca ttgtgaaaat ccttgatgaa 3900
aatgacaaca aacctcagtt tctgcaaaag ttctacaaaa tcagactccc tgagcgggaa 3960
aagccagacc gagaaagaaa tgccagacgg gagccgctct atcgcgtcat agccaccgac 4020
aaggatgagg gccccaatgc agaaatctcc tacagcatcg aagacgggaa tgagcatggc 4080
aaatttttca tcgaaccgaa aactggagtg gtttcgtcca agaggtttc agcagctgga 4140
gaatatgata ttctttcaat taaggcagtt gacaatggtc gccctcaaaa gtcatcaacc 4200
accagactcc atattgaatg gatctccaag cccaaacagt ccctggagcc catttcattt 4260
gaagaatcat tttttacctt tactgtgatg gaaagtgacc ccgttgctca catgattgga 4320
gtaatatctg tggagcctcc tggcataccc ctttggtttg acatcactgg tggcaactac 4380
gacagtcact tcgatgtgga caagggaact ggaaccatca ttgttgccaa acctcttgat 4440
gcagaacaga agtcaaacta caacctcaca gtcgaggcta cagatggaac caccactatc 4500
ctcactcagg tattcatcaa agtaatagac acaaatgacc atcgtcctca gttttctaca 4560
tcaaagtatg aagttgttat tcctgaagat acagcgccag aaacagaaat tttgcaaatc 4620
agtgctgtgg atcaggatga gaaaaacaaa ctaatctaca ctctgcagag cagtagagat 4680
ccactgagtc tcaagaaatt tcgtcttgat cctgcaaccg gctctctcta tacttctgag 4740
aaactggatc atgaagctgt ttcaccagca cacctcacgg tcatggtacg agatcaagat 4800
gtgcctgtaa aacgcaactt tgcaaggatt gtggtcaatg tcagcgacac gaatgaccac 4860
gccccgtggt tcaccgcttc ctcctacaaa gggcgggttt atgaatcggc agccgttggc 4920
tcagttgtgt tgcaggtgac ggctctggac aaggacaaag ggaaaaatgc tgaagtgctg 4980
tactcgatcg agtcaggaaa tattggaaat attggaaatt cttttatgat tgatcctgtc 5040
ttgggctcta ttaaaactgc caaagaatta gatcgaagta accaagcgga gtatgattta 5100
atggtaaaag ctacagataa gggcagtcca ccaatgagtg aaataacttc tgtgcgtatc 5160
tttgtcacaa ttgctgacaa cgcctctccg aagtttacat caaaagaata ttctgttgaa 5220
cttagtgaaa ctgtcagcat tgggagtttc gttgggatgg ttacagccca tagtcaatca 5280

tcagtggtgt atgaaataaa agatggaaat acaggtgatg cttttgatat taatccacat 5340
tctggaacta tcatcactca gaaagccctg gactttgaaa ctttgcccat ttacacattg 5400
ataatacaag gaactaacat ggctggtttg tccactaata caacggttct agttcacttg 5460
caggatgaga atgacaacgc gccagttttt atgcaggcag aatatacagg actcattagt 5520
gaatcagcct caattaacag cgtggtccta acagacagga atgtcccact ggtgattcga 5580
gcagctgatg ctgataaaga ctcaaatgct ttgcttgtat atcacattgt tgaaccatct 5640
gtacacacat attttgctat tgattctagc actggtgcta ttcatacagt actaagtctg 5700
gactatgaag aaacaagtat ttttcacttt accgtccaag tgcatgacat gggaacccca 5760
cgtttatttg ctgagtatgc agcgaatgta acagtacatg taattgacat taatgactgc 5820
ccccctgtgt ttgccaagcc attatatgaa gcatctcttt tgttaccaac atacaaagga 5880
gtaaaagtca tcacagtaaa tgctacagat gctgattcaa gtgcattctc acagttgatt 5940
tactccatca ccgaaggcaa catcggggag aagttttcta tggactacaa gactggtgct 6000
ctcactgtcc aaaacacaac tcagttaaga agccgctacg agctaaccgt tagagcttcc 6060
gatggcagat ttgccggcct tacctctgtc aaaattaatg tgaaagaaag caaagaaagt 6120
cacctaaagt ttacccagga tgtctactct gcggtagtga aagagaattc caccgaggcc 6180
gaaacattag ctgtcattac tgctattggg agtccaatca atgagccttt gttttatcac 6240
atcctcaacc cagatcgcag atttaaaata agccgcactt caggggttct gtcaaccact 6300
ggcacgccct tcgatcgtga gcagcaggag gcgtttgatg tggttgtaga agtgatagag 6360
gaacataagc cttctgcagt ggcccacgtt gtcgtgaagg tcattgtaga agaccaaaat 6420
gataatgcgc cggtgtttgt caaccttccc tactacgccg ttgttaaagt ggacactgag 6480
gtgggccatg tcattcgcta tgtcactgct gtagacagag acagtggcag aaacggggaa 6540
gtgcattact acctcaagga acatcatgaa cactttcaaa ttggaccctt gggtgaaatt 6600
tcactgaaaa agcaatttga gcttgacacc ttaaataaag aatatcttgt tacagtggtt 6660
gcaaaagatg gagggaaccc ggccttttca gcggaagtta tcgttccgat cactgtcatg 6720
aataaagcca tgcctgtgtt tgaaaaacct ttctacagtg cagagattgc agagagcatc 6780
caggtgcaca gccctgtggt ccacgtgcag gctaacagcc cggaaggcct gaaagtgttc 6840
tacagcatca cagacggaga ccctttcagc cagttcacta ttaacttcaa tactggagtt 6900
atcaatgtca tagctcctct ggactttgag gcccacccgg catataagct gagcatacgc 6960
gcaactgact ccttgacggg cgctcatgct gaagtatttg tggacatcat agtagacgac 7020
atcaatgata accctcctgt gtttgctcag cagtcttatg cggtgaccct gtctgaggca 7080

tctgtaattg gaacgtctgt tgttcaagtt agagccaccg attctgattc agaaccaaat 7140

agaggaatct cataccagat gtttgggaat cacagcaaga gtcatgatca ttttcatgta 7200

gacagcagca ctggcctcat ctcactactc agaaccctgg attacgagca gtcccggcag 7260

cacacgattt ttgtgagggc agttgatggt ggtatgccca cgctgagcag tgatgtgatt 7320

gtcacggtgg acgttaccga cctcaatggt aatccaccac tctttgaaca acagatttat 7380

gaagccagaa ttagcgagca cgcccctcat gggcatttcg tgacctgtgt aaaagcctat 7440

gatgcagaca gttcagacat agacaagttg cagtattcca ttctgtctgg caatgatcat 7500

aaacattttg tcattgacag tgcaacaggg attatcaccc tctcaaacct gcaccggcac 7560

gccctgaagc cattttacag tcttaacctg tcagtgtctg atggagtttt tagaagttcc 7620

acccaggttc atgtaactgt aattggaggc aatttgcaca gtcctgcttt ccttcagaac 7680

gaatatgaag tggaactagc tgaaaacgct cccctacata ccctggtgat ggaggtgaaa 7740

actacggatg gggattctgg tatttatggt cacgttactt accatattgt aaatgacttt 7800

gccaaagaca gattttacat aaatgagaga ggacagatat ttactttgga aaaacttgat 7860

cgagaaaccc cggcggagaa agtgatctca gtccgtttaa tggctaagga tgctggagga 7920

aaagttgctt tctgcaccgt gaatgtcatc cttacagatg acaatgacaa tgcaccacaa 7980

tttcgagcaa ccaaatacga agtgaatatc gggtccagtg ctgctaaagg gacttcagtc 8040

gtaaagtctg caagtgatgc cgatgagggc tccaatgccg acatcaccta tgccattgaa 8100

gcagactctg aaagtgtaaa agagaatttg gaaattaaca aactgtccgg cgtaatcact 8160

acaaaggaga gcctcattgg cttggaaaat gaattcttca ctttctttgt tagagctgtg 8220

gataatgggt ctccatcaaa agaatctgtt gttcttgtct atgttaaaat ccttccaccg 8280

gaaatgcagc ttccaaaatt ttcagaacct ttctatacct ttacagtgtc agaggacgtg 8340

cctgttggaa cagagataga tctcatccga gcagaacata gtgggactgt tctttacagc 8400

ctggtcaaag ggaatactcc agaaagcaat agggatgagt cctttgtgat tgacagacag 8460

agcgggagac tgaagttgga gaagagtctt gatcatgaga caactaagtg gtatcagttt 8520

tccatactgg ccaggtgcac tcaagatgac catgagatgg tggcttctgt agatgttagt 8580

atccaagtga aagatgcaaa tgacaacagc ccggtctttg aatctagtcc atatgaggca 8640

ttcattgttg aaaacctgcc agggggaagt agagtaattc agatcagggc atctgatgct 8700

gactcaggaa ccaacggcca agttatgtat agcctggatc agtcacaaag tgtggaagtc 8760

attgaatcct ttgccattaa catggaaaca ggctggatta caactttaaa ggaacttgac 8820

catgaaaaga gagacaatta ccagattaaa gtggttgcat cagatcatgg tgaaaagatc 8880

cagctatcct ccacagccat tgtggatgtt accgtcaccg atgtcaacga tagtccacca 8940

cgattcacgg ccgagatcta taaagggact gtgagtgagg atgaccccca aggtggggtg 9000
attgccatct taagtaccac ggatgctgat tctgaagaga tcaacagaca agttacatat 9060
ttcataacag gaggggatcc tttaggacag tttgccgttg aaactataca gaatgaatgg 9120
aaggtatatg tgaagaaacc tctagacagg gaaaaaaggg acaattacct tcttactatc 9180
acggcaactg atggcacctt ctcatcaaaa gcgatagttg aagtgaaagt tctggatgca 9240
aatgacaaca gtccagtttg tgaaaagact ttatattcag acactattcc tgaagacgtc 9300
cttcctggaa aattgatcat gcagatctct gctacagacg cagacatccg ctctaacgct 9360
gaaattactt acacgttatt gggttcaggt gcagaaaaat tcaaactaaa tccagacaca 9420
ggtgaactga aaacgtcaac ccccttgat cgtgaggagc aagctgttta tcatcttctc 9480
gtcagggcca cagatggagg aggaagattc tgccaagcca gtattgtcgt cacgctagaa 9540
gatgtgaacg ataacgcccc cgaattctct gccgatcctt atgccatcac cgtgtttgaa 9600
aacacagagc cgggaacgct gctgacaaga gtgcaggcca cagatgccga cgcaggatta 9660
aatcggaaga ttttatactc actgattgac tctgctgatg ggcagttctc cattaacgaa 9720
ttatctggaa ttattcagtt agaaaaacct ttggacagag aactccaggc agtatacacc 9780
ctctctttga aagctgtgga tcaaggcttg ccaaggaggc tgactgccac tggcactgtg 9840
attgtatcag ttcttgacat aaatgacaac ccccctgtgt ttgagtaccg tgaatatggt 9900
gccaccgtgt ctgaggacat tcttgttgga actgaagttc ttcaagtgta tgcagcaagt 9960
cgggatattg aagcaaatgc agaaatcacc tactcaataa taagtggaaa tgaacatggg 10020
aaattcagca tagattctaa aacaggggcc gtatttatca ttgagaatct ggattatgag 10080
agctctcatg agtattacct aacagtagag gccactgatg gaggcacgcc ttcactgagc 10140
gacgttgcca ctgtgaacgt taatgtaaca gatatcaacg ataatacccc tgtgttcagc 10200
caagacacct acacgacagt catcagtgaa gatgccgttc ttgagcagtc tgtcatcacg 10260
gttatggccg atgatgccga tggaccttcc aacagccaca tccactactc aattatagat 10320
ggcaaccaag gaagctcgtt cacaattgac cccgtcaggg gagaagtcaa agtgaccaaa 10380
cttctcgacc gagaaacgat ttcaggttac acgctcacgg ttcaagcttc tgataatggc 10440
agtccaccca gagtcaacac gacgaccgtg aacatcgatg tgtccgatgt caatgacaac 10500
gcgcccgtct tctccagggg aaactacagt gtcattatcc aggaaaataa gccagtgggc 10560
ttcagcgtgc tgcagctggt agtaacagat gaggattctt cccataacgg tccacccttc 10620
ttctttacta ttgtaactgg aaatgatgag aaggcttttg aagttaaccc gcaaggagtc 10680
ctcctgacat catctgccat caagaggaag gagaaagatc attacttact gcaggtgaag 10740

gtggcagata atggaaagcc tcagttgtca tctttgacat acattgacat tagggtaatt 10800
gaggagagca tctatccgcc tgcgattttg cccctggaga ttttcatcac ctcttctgga 10860
gaagaatact caggtggcgt cattgggaag atccatgcca cagaccagga cgtgtatgat 10920
actctaacct acagtctcga ccctcagatg gacaacctgt tctctgtttc cagcacaggg 10980
ggcaagctga tagcacacaa aaagctagac atagggcaat accttctcaa tgtcagcgta 11040
acagatggga agttcacgac ggtggccgac atcacagtgc atatcagaca agtcacacag 11100
gagatgttga accacaccat cgcgatccgc tttgccaacc tcactccgga agaattcgtt 11160
ggtgactact ggcgcaactt ccagcgagct ttacggaaca tcctgggtgt gaggaggaac 11220
gacatacaga ttgttagttt gcagtcctct gaacctcacc cacatctgga cgtcttactt 11280
tttgtagaga aaccaggtag tgctcagatc tcaacaaaac aacttctgca caagattaac 11340
tcttccgtga ctgacattga ggaaatcatt ggagttagga tactgaatgt attccagaaa 11400
ctctgcgcgg gactggactg cccctggaag ttctgcgatg aaaaggtgtc tgtggatgaa 11460
agtgtgatgt caacacacag cacagccaga ctgagttttg tgactccccg ccaccacagg 11520
gcagcggtgt gtctctgcaa agagggaagg tgcccacctg tccaccatgg ctgtgaagat 11580
gatccgtgcc ctgagggatc cgaatgtgtg tctgatccct gggaggagaa acacacctgt 11640
gtctgtccca gcggcaggtt tggtcagtgc ccagggagtt catctatgac actgactgga 11700
aacagctacg tgaaataccg tctgacggaa aatgaaaaca aattagagat gaaactgacc 11760
atgaggctca gaacatattc cacgcatgcg gttgtcatgt atgctcgagg aactgactat 11820
agcatcttgg agattcatca tggaaggctg cagtacaagt ttgactgtgg aagtggccct 11880
ggaattgtct ctgttcagag cattcaggtc aatgatgggc agtggcacgc agtggccctg 11940
gaagtgaatg gaaactatgc tcgcttggtt ctagaccaag ttcatactgc atcgggcaca 12000
gccccaggga ctctgaaaac cctgaacctg gataactatg tgtttttggg tggccacatc 12060
cgtcagcagg gaacaaggca tggaagaagt cctcaagttg gtaatggttt caggggttgt 12120
atggactcca tttatttgaa tgggcaggag ctccctttaa acagcaaacc cagaagctat 12180
gcacacatcg aagagtcggt ggatgtatct ccaggctgct tcctgacggc cacggaagac 12240
tgcgccagca acccttgcca gaatggaggc gtttgcaatc cgtcacctgc tggaggttat 12300
tactgcaaat gcagtgcctt gtacataggg acccactgtg agataagcgt caatccgtgt 12360
tcctccaacc catgcctcta tgggggcacg tgtgttgtcg acaacggagg ctttgtttgc 12420
cagtgtagag gattatatac tggtcagagg tgtcagctta gtccatactg caaagatgaa 12480
ccctgtaaga atggcggaac atgctttgac agtttggatg gcgccgtttg tcagtgtgat 12540
tcgggtttta ggggagaaag gtgtcagagt gatatcgacg agtgctctgg aaacccttgc 12600

ctgcacgggg ccctctgtga gaacacgcac ggctcctatc actgcaactg cagccacgag 12660
tacaggggac gtcactgcga ggatgctgcg cccaaccagt atgtgtccac gccgtggaac 12720
attgggttgg cggaaggaat tggaatcgtt gtgtttgttg cagggatatt tttactggtg 12780
gtggtgtttg ttctctgccg taagatgatt agtcggaaaa agaagcatca ggctgaacct 12840
aaagacaagc acctgggacc cgctacggct ttcttgcaaa gaccgtattt tgattccaag 12900
ctaaataaga acatttactc agacatacca ccccaggtgc ctgtccggcc tatttcctac 12960
accccgagta ttccaagtga ctcaagaaac aatctggacc gaaattcctt cgaaggatct 13020
gctatcccag agcatcccga attcagcact tttaaccccg agtctgtgca cgggcaccga 13080
aaagcagtgg cggtctgcag cgtggcgcca aacctgcctc ccccaccccc ttcaaactcc 13140
ccttctgaca gcgactccat ccagaagcct agctgggact ttgactatga cacaaaagtg 13200
gtggatcttg atccctgtct ttccaagaag cctctagagg aaaagccttc ccagccatac 13260
agtgcccggg aaagcctgtc tgaagtgcag tccctgagct ccttccagtc cgaatcgtgc 13320
gatgacaatg ggtatcactg ggatacatca gattggatgc caagcgttcc tctgccggac 13380
atacaagagt tccccaacta tgaggtgatt gatgagcaga cacccctgta ctcagcagat 13440
ccaaacgcca tcgatacgga ctattaccct ggaggctacg acatcgaaag tgattttcct 13500
ccacccccag aagacttccc cgcagctgat gagctaccac cgttaccgcc cgaattcagc 13560
aatcagtttg aatccatcca ccctcctaga gacatgcctg ccgcgggtag cttgggttct 13620
tcatcaagaa accggcagag gttcaacttg aatcagtatt tgcccaattt ttatcccctc 13680
gatatgtctg aacctcaaac aaaaggcact ggtgagaata gtacttgtag agaaccccat 13740
gcccccttacc cgccagggta tcaaagacac ttcgaggcgc ccgctgtcga gagcatgccc 13800
atgtctgtgt acgcctccac cgcctcctgc tctgacgtgt cagcctgctg cgaagtggag 13860
tccgaggtca tgatgagtga ctatgagagc ggggacgacg gccacttcga agaggtgacg 13920
atcccgcccc tggattccca gcagcacacg gaagtctgac tctcaactcc ccccaaagtg 13980
cctgacttta gtgaacctag aggtgatgtg agtaatccgc gctgttcttt gcagcagtgc 14040
ttccaagctt tttttggtga gccgaatggg catggctgcg ctggatcctg cgcctctgga 14100
cgtgctagcc atttccagtg tcccaactac tgtcatcgtg aggttttcat cggctgtgcc 14160
atttcccaac gtcttttggg atttacatct gtctgtgtta aaataatcaa acgaaaaatc 14220
agtcctgtgt tgtcagcatg attcatgtat ttatatagat ttgattattt taattttcct 14280
gtctcttttt tttgtaaatt ttatgtacag atttgatttt tcatagtttt aactagattt 14340
ccaagatatt ttgtgcattt gtttcaactg aattttggtg gtgtcagtgc cattatctag 14400

caccctgatt ttttttttt tactataacc agggtttcat tctgtctttt tccactgaag 14460

tgtgacattt tgttagtaca tttcagtgta gtcattcatt tctagctgta cataggatga 14520

aggagagatc agatacatga acatgtctta catgggttgc tgtatttaga attataaaca 14580

tttttcatta ttggaaagtg taacggggac cttctgcata cctgtttaga accaaaacca 14640

ccatgacaca gtttttatag tgtctgtata tttgtgatgc aatggtcttg taaaggtttt 14700

taatgaaaac taccattagc cagtctttct tactgacaat aaattattaa taaaat 14756

<210> 36
<211> 1583
<212> DNA
<213> Homo Sapiens

<400> 36
gtggtgtttg ctttctccac cagaagggca cactttcatc taatttgggg tatcactgag 60

ctgaagacaa agagaagggg gagaaaacct agcagaccac catgtgctat gggaagtgtg 120

cacgatgcat cggacattct ctggtggggc tcgccctcct gtgcatcgcg gctaatattt 180

tgctttactt tcccaatggg gaaacaaagt atgcctccga aaaccacctc agccgcttcg 240

tgtggttctt ttctggcatc gtaggaggtg gcctgctgat gctcctgcca gcatttgtct 300

tcattgggct ggaacaggat gactgctgtg gctgctgtgg ccatgaaaac tgtggcaaac 360

gatgtgcgat gctttcttct gtattggctg ctctcattgg aattgcagga tctggctact 420

gtgtcattgt ggcagccctt ggcttagcag aaggaccact atgtcttgat tccctcggcc 480

agtggaacta caccttgcc agcaccgagg gccagtacct tctggatacc tccacatggt 540

ccgagtgcac tgaacccaag cacattgtgg aatggaatgt atctctgttt tctatcctct 600

tggctcttgg tggaattgaa ttcatcttgt gtcttattca agtaataaat ggagtgcttg 660

gaggcatatg tggcttttgc tgctctcacc aacagcaata tgactgctaa aagaaccaac 720

ccaggacaga gccacaatct tcctctattt cattgtaatt tatatatttc acttgtattc 780

atttgtaaaa ctttgtatta gtgtaacata ctccccacag tctactttta caaacgcctg 840

taaagactgg catcttcaca ggatgtcagt gtttaaattt agtaaacttc ttttttgttt 900

gtttatttgt ttttgttttt ttttaaggaa tgaggaaaca aaccaccctc tgggggtagt 960

ttacagactg agtgacagta ctcagtatat ctgagataaa ctctataatg ttttggataa 1020

aaataacatt ccaatcacta ttgtatatat gtgcatgtat tttttaaatt aaagatgtct 1080

agttgctttt tataagacca agaaggagaa aatccgacaa cctggaaaga tttttgtttt 1140

cactgcttgt atgatgtttc ccattcatac acctataaat ctctaacaag aggccctttg 1200

aactgccttg tgttctgtga gaaacaaata tttacttaga gtggaaggac tgattgagaa 1260

tgttccaatc caaatgaatg catcacaact tacaatgctg ctcattgttg tgagtactat 1320

gagattcaaa tttttctaac atatggaaag ccttttgtcc tccaaagatg agtactaggg 1380

atcatgtgtt taaaaaaaga aaggctacga tgactgggca agaagaaaga tgggaaactg 1440

aataaagcag ttgatcagca tcattggaac atggggacga gtgacggcag gaggaccacg 1500

aggaaatacc ctcaaaacta acttgtttac aacaaaataa agtattcact acgaaaaaaa 1560

aaaaaaaaaa aaaaaaaaaa aaa 1583

<210> 37
<211> 7586
<212> DNA
<213> Homo Sapiens

<400> 37
gttctttgtg acacatcaca cagaattgga gtgctgtcct tctggagagt ggtggagaac 60

caagatacag ttcagaacca aaggaataga gaagggcttt gatttctttt tggctttaga 120

ttggggattt gggaggctta gcaggaaaga tgtccactga aaatgtggaa gggaagccca 180

gtaaccttgg ggagagagga agagcccgga gctccacttt cctcagggtt gtccagccaa 240

tgtttaacca cagtattttc acttctgcag tctctcctgc tgcagaacgc atccgattca 300

tcttgggaga ggaggatgac agcccagctc cccctcagct cttcacggaa ctggatgagc 360

tgctggccgt ggatgggcag gagatggagt ggaaggaaac agccaggtgg atcaagtttg 420

aagaaaaagt ggaacagggt ggggaaagat ggagcaagcc ccatgtggcc acattgtccc 480

ttcatagttt atttgagctg aggacatgta tggagaaagg atccatcatg cttgatcggg 540

aggcttcttc tctcccacag ttggtggaga tgattgttga ccatcagatt gagacaggcc 600

tattgaaacc tgaacttaag gataaggtga cctatacttt gctccggaag caccggcatc 660

aaaccaagaa atccaacctt cggtccctgg ctgacattgg gaagacagtc tccagtgcaa 720

gtaggatgtt taccaaccct gataatggta gcccagccat gacccatagg aatctgactt 780

cctccagtct gaatgacatt tctgataaac cggagaagga ccagctgaag aataagttca 840

tgaaaaaatt gccacgtgat gcagaagctt ccaacgtgct tgttggggag gttgactttt 900

tggatactcc tttcattgcc tttgttaggc tacagcaggc tgtcatgctg ggtgccctga 960

ctgaagttcc tgtgcccaca aggttcttgt tcattctctt aggtcctaag gggaaagcca 1020

agtcctacca cgagattggc agagccattg ccaccctgat gtctgatgag gtgttccatg 1080

acattgctta taaagcaaaa gacaggcacg acctgattgc tggtatggat gagttcctag 1140

atgaagtcat cgtccttcca cctggggaat gggatccagc aattaggata gagcctccta 1200

agagtcttcc atcctctgac aaaagaaaga atatgtactc aggtggagag aatgttcaga 1260

tgaatgggga tacgccccat gatggaggtc acggaggagg aggacatggg gattgtgaag 1320

aattgcagcg aactggacgg ttctgtggtg gactaattaa agacataaag aggaaagcgc 1380

catttttgc cagtgatttt tatgatgctt taaatattca agctctttcg gcaattctct 1440

tcatttatct ggcaactgta actaatgcta tcacttttgg aggactgctt ggggatgcca 1500

ctgacaacat gcagggcgtg ttggagagtt tcctgggcac tgctgtctct ggagccatct 1560

tttgcctttt tgctggtcaa ccactcacta ttctgagcag caccggacct gtcctagttt 1620

ttgagaggct tctatttaat ttcagcaagg acaataattt tgactatttg gagtttcgcc 1680

tttggattgg cctgtggtcc gccttcctat gtctcatttt ggtagccact gatgccagct 1740

tcttggttca atacttcaca cgtttcacgg aggagggctt ttcctctctg attagcttca 1800

tctttatcta tgatgctttc aagaagatga tcaagcttgc agattactac cccatcaact 1860

ccaacttcaa agtgggctac aacactctct tttcctgtac ctgtgtgcca cctgacccag 1920

ctaatatctc aatatctaat gacaccacac tggccccaga gtatttgcca actatgtctt 1980

ctactgacat gtaccataat actacctttg actgggcatt tttgtcgaag aaggagtgtt 2040

caaaatacgg aggaaacctt gtcgggaaca actgtaattt tgttcctgat atcacactca 2100

tgtcttttat cctcttcttg ggaacctaca cctcttccat ggctctgaaa aaattcaaaa 2160

ctagtcctta ttttccaacc acagcaagaa aactgatcag tgattttgcc attatcttgt 2220

ccattctcat cttttgtgta atagatgccc tagtaggcgt ggacacccca aaactaattg 2280

tgccaagtga gttcaagcca acaagtccaa accgaggttg gttcgttcca ccgtttggag 2340

aaaaccctg gtgggtgtgc cttgctgctg ctatcccggc tttgttggtc actatactga 2400

ttttcatgga ccaacaaatt acagctgtga ttgtaaacag gaaagaacat aaactcaaga 2460

aaggagcagg gtatcacttg gatctctttt gggtggccat cctcatggtt atatgctccc 2520

tcatggctct tccgtggtat gtagctgcta cggtcatctc cattgctcac atcgacagtt 2580

tgaagatgga gacagagact tctgcacctg gagaacaacc aaagtttcta ggagtgaggg 2640

aacaaagagt cactggaacc cttgtgttta ttctgactgg tctgtcagtc tttatggctc 2700

ccatcttgaa gtttataccc atgcctgtac tctatggtgt gttcctgtat atgggagtag 2760

catcccttaa tggtgtgcag ttcatggatc gtctgaagct gcttctgatg cctctgaagc 2820

atcagcctga cttcatctac ctgcgtcatg ttcctctgcg cagagtccac ctgttcactt 2880

tcctgcaggt gttgtgtctg gccctgcttt ggatcctcaa gtcaacggtg gctgctatca 2940

tttttccagt aatgatcttg gcacttgtag ctgtcagaaa aggcatggac tacctcttct 3000

cccagcatga cctcagcttc ctggatgatg tcattccaga aaaggacaag aaaaagaagg 3060

aggatgagaa gaaaaagaaa aagaagaagg gaagtctgga cagtgacaat gatgattctg 3120

actgcccata ctcagaaaaa gttccaagta ttaaaattcc aatggacatc atggaacagc 3180
aacctttcct aagcgatagc aaaccttctg acagagaaag atcaccaaca ttccttgaac 3240
gccacacatc atgctgataa aattcctttc cttcagtcac tcggtatgcc aagtcctcct 3300
agaactccag taaaagttgc ctcaaattag actagaactt gaacctgaag acaatgatta 3360
tttctggagg agcaagggaa cagaaactac attgtaacct gtttgtcttt cttaaaactg 3420
acatttgttg ttaatgtcat ttgtttttgt ttggctgttt gtttattttt taacttttat 3480
ttcgtctcag tttttggtca caggccaaat aatacagcgc tctctctgct tctctcttgc 3540
atagatacaa tcaagacaat agtgcaccgt tccttaaaaa cagcatctga ggaatccccc 3600
ttttgttctt aaactttcag atgtgtcctt tgataaccaa attctgtcac tcaagacaca 3660
gacacccaca gaccctgtcc tttgcctcta ttaagcagag gatggaagta ttaaggattt 3720
tgtaacacct tttatgaaaa tgttgaagga acttaaaact ttagctttgg agctgtgctt 3780
actggcttgt ctttgtctgg tagaacaaac cttgacctcc agacagagtc ccttctcact 3840
tatagagctc tccaggactg gaaaaagtgc tgctatttta acttgctctt gcttgtaaat 3900
cctaatctta gagttatcaa aagaagaaaa aactgaaggt actttactcc ctatagagaa 3960
accattgcca tcattgtagc aagtgctgga atgtcccttt tttcctatgc aacttttta 4020
taacccttta atgaacttat ctgtggagta cattgaagaa tatttttctt cctagatttt 4080
gttgtttaaa ttatggggcc taacctgcca cttatttttt gtcaattttt aaaacttttt 4140
tttaattact gtaaagaaaa tgaatttttt cctgcagcag gaaacatagt tttcagtagt 4200
tctacctctt atttgtagct gccaggcttt ctgtaaaaat tgtattgtat ataatgtgat 4260
ttttacacat acatacacac acaaatacac aatctctagg gtaagccaga aggcaagatc 4320
agattaaaaa caccatgttt ctaagcatcc atttttccct ttctttaaaa gaaacttaac 4380
tgttctatga aggagattga gggagaagag acaaactcct atgtcatgag aataaccgat 4440
gttctgataa tagtagcatc taggtacaga tgctggttgt attaccacgt caatgtccta 4500
tgcagtattg ttagacattt tctcattttg aaatatttgt gtgtttgtgt atgtgctctg 4560
tgccatggct ggtgtatata tgtgcaatgt tagaaggcaa aagagtgatg gtaggcagag 4620
ggcaaagtca ttgaatctct tatgccagtt ttcataaaac ccaaaccaca tatgaaaaaa 4680
tccattaagg gtccaagaag tctgtccata tgaaaatgag ggtaaatata gtttatttcc 4740
caggtatcag tcattataat tgatataata gctctaacat gcaatataaa attcatagga 4800
gtattaatag cccatttaca catctataaa atgtaatggg attgcagagc tgcagagtac 4860
agtgtaacag tactctcatg caatttttt caggatgcaa aggcaattat tctttgtaag 4920

cgggacattt agatatattt gtgtacatat tatatgtatg tatatttcaa agtaccacac 4980
tgaaaattag acatttatta accaaattta acgtggtatt taaaggtaat atttttaata 5040
tgatacatta catattgtga atgtatacta aaaaaacatt ttaaatgtta aaattataat 5100
ttcagattca tataaccaca actgtgatat atcctaacta taaccagttg ttgaggggta 5160
tactagaagc agaatgaaac cacatttttt ggtttgataa tatgcactta ttgactccca 5220
ctcattgtta tgttaattaa gttattattc tgtctccttg taattttgat tacaaaaatt 5280
ttattatcct gagttagctg ttacttttac agtacctgat actcctaaaa cttttaactt 5340
atacaaatta gtcaataatg accccaattt tttcattaaa ataatagtgg tgaattatat 5400
gttattgtgt taaaacctca cttgccaaat tctggcttca catttgtatt tagggctatc 5460
cttaaaatga tgagtctata ttatctagct ttctattacc ctaatataaa ctggtataag 5520
aagactttcc ttttttcttt atgcatggaa gcatcaataa attgtttaaa aaccatgtat 5580
agtaaattca gcttaacccg tgatcttctt aagttaaagg tacttttgtt ttataaaagc 5640
tctagataaa actttctttt ctgatcatga atcaagtatc tgtggtttca tgcccctctc 5700
tatacctttc aaagaactcc tgaagcaact taactcatca tttcagcctc tgagtagagg 5760
taaaacctat gtgtacttct gtttatgatc catattgata tttatgacat gaacacagaa 5820
tagtaccttа catttgctaa acagacagtt aatatcaaat cctttcaata ttctgggaac 5880
ccagggaagt ttttaaaaat gtcattactt tcaaaggaac agaagtagtt aaccaaacta 5940
acaagcaaaa cctgaggttt acctagtgac accaaattat cggtatttta actgaattta 6000
cccattgact aagaatgaac cggatttggt ggtggttttg tttctatgca aactggacac 6060
aaattacaac agtaaatttt tttataagtg cttctccctt ctccatgatg tgacttccgg 6120
agataaagga ttcaaaagat aaagacaaag tacgctcaga gttgttaacc agaaagtcct 6180
ggctgtggtt gcagaaacac tgttggaaga aaagagatga ctaagtcaag tgtctgcctt 6240
atcaaaagag caaaaatgcc tctggttttg tgtttgggag aaaaatatct tggacgcact 6300
gttttccttg ataaaagtca tcttctctac tgtgtgaaat gaatacttgg aattctaatt 6360
gttttgtgtg ccaggggcag taatgtccct gcctcttctc ccaatcaagg ttgaggagtg 6420
gggctgggga gaggacttaa ctgacttaag aagtaggaaa acaaaaacct ctctcctcag 6480
ccttccacct ccaagagagg aggaaaaaca gttgtctgct gtctgtaatt cagtttgcgt 6540
gtattttatg ctcatgcacc aacccataca gagtaaatct tttatcaact atatactggt 6600
gtttaataga gaatgattgt cttccgagtt ttttggttcc ttttttaact gtgttaaagt 6660
acttgaaatg tattgactgc tgactatatt ttaaaaacaa aatgaaataa tttgagttgt 6720
attacagagg ttgacattgt tcagggatgg gacaaagcct tcttcaatcc ttttcatact 6780

acttaatgat tttggtgcag gaacctgaga ttttctgatt tatatttcat gatatttcac 6840

atttgctctt cacagcatga gcatgaagcc cagtggcacc aaatggctgg gtacaatcaa 6900

gtgatatttt gtagcacctc actatctgaa aggccatgag ttttcagatg atttcattga 6960

gcttcattgc agcctgaaat tttaaaaaag ttgtgtaata cgccaaccag tcaagttgtg 7020

ttttggccag agatttagat atgtccaatt tcctggctca tttcattgtg ctctatgggt 7080

acgtataaaa agcaagaatt ctgtttccta ggcaaacatt gcaactcagg gctaaagtca 7140

tccagtgaaa cttttagagc cagaagtaac tttgtcccag tcctacaatg tgaaaagagt 7200

gaatagttgc ctctttttag ccattttcat ggctggtaca tattcgtacg cattactttt 7260

cagaatcaat acgcactttc agatattctt atttttattc tcttaagtct ttattaactt 7320

tggagagaga aatgatgcat ctttttattt taaatgaagt agatcaacat ggtggaacaa 7380

aatgataaag aacagaaaac atttcaatat attactaata acttttttcca atataaatcc 7440

taaaattcct ataacatagt attttacagt tttatgaagc tttctattgt gacttttatg 7500

gaattaagag atgaagaaga tgagatattt tagcatttat atttttcaaa attatatgta 7560

tacttaaaaa taaagtaact ttatgc 7586

<210> 38
<211> 1958
<212> DNA
<213> Homo Sapiens

<400> 38
cggcacagcc tcacacctga acgctgtcct cccgcagacg agaccggcgg gcactgcaaa 60

gctgggactc gtctttgaag gaaaaaaaat agcgagtaag aaatccagca ccattcttca 120

ctgacccatc ccgctgcacc tcttgtttcc caagtttttg aaagctggca actctgacct 180

cggtgtccaa aaatcgacag ccactgagac cggctttgag aagccgaaga tttggcagtt 240

tccagactga gcaggacaag gtgaaagcag gttggaggcg ggtccaggac atctgagggc 300

tgaccctggg ggctcgtgag gctgccaccg ctgctgccgc tacagaccca gccttgcact 360

ccaaggctgc gcaccgccag ccactatcat gtccactccc ggggtcaatt cgtccgcctc 420

cttgagcccc gaccggctga acagcccagt gaccatcccg gcggtgatgt tcatcttcgg 480

ggtggtgggc aacctggtgg ccatcgtggt gctgtgcaag tcgcgcaagg agcagaagga 540

gacgaccttc tacacgctgg tatgtgggct ggctgtcacc gacctgttgg gcactttgtt 600

ggtgagcccg gtgaccatcg ccacgtacat gaagggccaa tggcccgggg gccagccgct 660

gtgcgagtac agcaccttca ttctgctctt cttcagcctg tccggcctca gcatcatctg 720

cgccatgagt gtcgagcgct acctggccat caaccatgcc tatttctaca gccactacgt 780

ggacaagcga ttggcgggcc tcacgctctt tgcagtctat gcgtccaacg tgctcttttg 840

cgcgctgccc aacatgggtc tcggtagctc gcggctgcag tacccagaca cctggtgctt 900

catcgactgg accaccaacg tgacggcgca cgccgcctac tcctacatgt acgcgggctt 960

cagctccttc ctcattctcg ccaccgtcct ctgcaacgtg cttgtgtgcg gcgcgctgct 1020

ccgcatgcac cgccagttca tgcgccgcac ctcgctgggc accgagcagc accacgcggc 1080

cgcggccgcc tcggttgcct cccggggcca ccccgctgcc tccccagcct tgccgcgcct 1140

cagcgacttt cggcgccgcc ggagcttccg ccgcatcgcg ggcgccgaga tccagatggt 1200

catcttactc attgccacct ccctggtggt gctcatctgc tccatcccgc tcgtggtgcg 1260

agtattcgtc aaccagttat atcagccaag tttggagcga gaagtcagta aaaatccaga 1320

tttgcaggcc atccgaattg cttctgtgaa ccccatccta gacccctgga tatatatcct 1380

cctgagaaag acagtgctca gtaaagcaat agagaagatc aaatgcctct tctgccgcat 1440

tggcgggtcc cgcagggagc gctccggaca gcactgctca gacagtcaaa ggacatcttc 1500

tgccatgtca ggccactctc gctccttcat ctcccgggag ctgaaggaga tcagcagtac 1560

atctcagacc ctcctgccag acctctcact gccagacctc agtgaaaatg gccttggagg 1620

caggaatttg cttccaggtg tgcctggcat gggcctggcc caggaagaca ccacctcact 1680

gaggactttg cgaatatcag agacctcaga ctcttcacag ggtcaggact cagagagtgt 1740

cttactggtg gatgaggctg gtgggagcgg cagggctggg cctgccccta aggggagctc 1800

cctgcaagtc acatttccca gtgaaacact gaacttatca gaaaaatgta tataataggc 1860

aaggaaagaa atacagtact gtttctggac ccttataaaa tcctgtgcaa tagacacata 1920

catgtcacat ttagctgtgc tcagaagggc tatcatca 1958

<210> 39
<211> 1740
<212> DNA
<213> Homo Sapiens

<400> 39
cagtatccct cctgacaaaa ctaacaaaaa tcctgttagc caaataatca gccacattca 60

tatttaccgt caaagttttt atcctcattt tacagcagtg gagagcgatt gccccgggtc 120

ccacgttagg aagagagaga actgggattt gcacccaggc aatctgggga cagagctgtg 180

atcacaactc catgagtcag ggccgagcca gcccccttcac caccagccgg ccgcgccccg 240

ggaaggaagt ttgtggcgga ggaggttcgt acgggaggag ggggaggcgc ccacgcatct 300

ggggctgact cgctctttcg caaaacgtct gggaggagtc cctggggcca caaaactgcc 360

tccttcctga ggccagaagg agagaagacg tgcagggacc ccgcgcacag gagctgccct 420

cgcgacatgg gtcacccgcc gctgctgccg ctgctgctgc tgctccacac ctgcgtccca 480

gcctcttggg gcctgcggtg catgcagtgt aagaccaacg gggattgccg tgtggaagag 540

tgcgccctgg gacaggacct ctgcaggacc acgatcgtgc gcttgtggga agaaggagaa 600

gagctggagc tggtggagaa aagctgtacc cactcagaga agaccaacag gaccctgagc 660

tatcggactg gcttgaagat caccagcctt accgaggttg tgtgtgggtt agacttgtgc 720

aaccagggca actctggccg ggctgtcacc tattcccgaa gccgttacct cgaatgcatt 780

tcctgtggct catcagacat gagctgtgag aggggccggc accagagcct gcagtgccgc 840

agccctgaag aacagtgcct ggatgtggtg acccactgga tccaggaagg tgaagaaggg 900

cgtccaaagg atgaccgcca cctccgtggc tgtggctacc ttcccggctg cccgggctcc 960

aatggtttcc acaacaacga caccttccac ttcctgaaat gctgcaacac caccaaatgc 1020

aacgagggcc caatcctgga gcttgaaaat ctgccgcaga atggccgcca gtgttacagc 1080

tgcaagggga acagcaccca tggatgctcc tctgaagaga ctttcctcat tgactgccga 1140

ggccccatga atcaatgtct ggtagccacc ggcactcacg aaccgaaaaa ccaaagctat 1200

atggtaagag gctgtgcaac cgcctcaatg tgccaacatg cccacctggg tgacgccttc 1260

agcatgaacc acattgatgt ctcctgctgt actaaaagtg gctgtaacca cccagacctg 1320

gatgtccagt accgcagtgg ggctgctcct cagcctggcc ctgcccatct cagcctcacc 1380

atcaccctgc taatgactgc cagactgtgg ggaggcactc tcctctggac ctaaacctga 1440

aatcccccctc tctgccctgg ctggatccgg gggacccctt tgcccttccc tcggctccca 1500

gccctacaga cttgctgtgt gacctcaggc cagtgtgccg acctctctgg gcctcagttt 1560

tcccagctat gaaaacagct atctcacaaa gttgtgtgaa gcagaagaga aaagctggag 1620

gaaggccgtg ggcaatggga gagctcttgt tattattaat attgttgccg ctgttgtgtt 1680

gttgttatta attaatattc atattattta ttttatactt acataaagat tttgtaccag 1740

<210> 40
<211> 3088
<212> DNA
<213> Homo Sapiens

<400> 40
ttgcttgagt catcttctga agctttaaaa acaattgatg aattggcctt caagatagac 60

ctaaatagca catcacatgt gaatattaca actcggaact tggctctcag cgtatcatcc 120

ctgttaccag ggacaaatgc aatttcaaat tttagcattg gtcttccaag caataatgaa 180

tcgtatttcc agatggattt tgagagtgga caagtggatc cactggcatc tgtaattttg 240

cctccaaact tacttgagaa tttaagtcca gaagattctg tattagttag aagagcacag 300

tttactttct tcaacaaaac tggacttttc caggatgtag gaccccaaag aaaaacttta 360
gtgagttatg tgatggcgtg cagtattgga aacattacta tccagaatct gaaggatcct 420
gttcaaataa aaatcaaaca tacaagaact caggaagtgc atcatcccat ctgtgccttc 480
tgggatctga acaaaaacaa aagttttgga ggatggaaca cgtcaggatg tgttgcacac 540
agagattcag atgcaagtga gacagtctgc ctgtgtaacc acttcacaca ctttggagtt 600
ctgatggacc ttccaagaag tgcctcacag ttagatgcaa gaaacactaa agtcctcact 660
ttcatcagct atattgggtg tggaatatct gctatttttt cagcagcaac tctcctgaca 720
tatgttgctt ttgagaaatt gcgaagggat tatccctcca aaatcttgat gaacctgagc 780
acagccctgc tgttcctgaa tctcctcttc ctcctagatg gctggatcac ctccttcaat 840
gtggatggac tttgcattgc tgttgcagtc ctgttgcatt tcttccttct ggcaaccttt 900
acctggatgg ggctagaagc aattcacatg tacattgctc tagttaaagt atttaacact 960
tacattcgcc gatacattct aaaattctgc atcattggct ggggtttgcc tgccttagtg 1020
gtgtcagttg ttctagcgag cagaaacaac aatgaagtct atggaaaaga aagttatggg 1080
aaagaaaaag gtgatgaatt ctgttggatt caagatccag tcatatttta tgtgacctgt 1140
gctgggtatt ttggagtcat gtttttttctg aacattgcca tgttcattgt ggtaatggtg 1200
cagatctgtg ggaggaatgg caagagaagc aaccggaccc tgagagaaga agtgttaagg 1260
aacctgcgca gtgtggttag cttgaccttt ctgttgggca tgacatgggg ttttgcattc 1320
tttgcctggg gacccttaaa tatccccttc atgtacctct tctccatctt caattcatta 1380
caaggcttat ttatattcat cttccactgt gctatgaagg agaatgttca gaaacagtgg 1440
cggcggcatc tctgctgtgg tagatttcgg ttagcagata actcagattg gagtaagaca 1500
gctaccaata tcatcaagaa aagttctgat aatctaggaa aatctttgtc ttcaagctcc 1560
attggttcca actcaaccta tcttacatcc aaatctaaat ccagctctac cacctatttc 1620
aaaaggaata gccacacaga taatgtctcc tatgagcatt ccttcaacaa aagtggatca 1680
ctcagacagt gcttccatgg acaagtcctt gtcaaaactg gcccatgctg atggagatca 1740
aacatcaatc atccctgtcc atcaggtcat tgataaggtc aagggttatt gcaatgctca 1800
ttcagacaac ttctataaaa atattatcat gtcagacacc ttcagccaca gcacaaagtt 1860
ttaatgtctt taagaaaaag aaatcaatct gcagaaatgt gaagatttgc aagcagtgta 1920
aactgcaact agtgatgtaa atgtgctatt acctaggtaa ctgcatatat ataaggaatg 1980
tattttgtta agaaggcttt tgtgaaattc agaatttttc tttttaatat atttcttcca 2040
tggaagagtt gtcatcacta aaacttcagt actgagagta acatgactca gtagccacag 2100

aagctatgat ttgtaaaata tataattgaa tcagagtaat cataatgcag gggagacatt 2160

caaattagag acaagggaga agcaatgctg aggaagaccc tagatagagc tcattttact 2220

ccacctaatc gttatatctg gatataccca ttttctgcat cttctttctc aacaataaac 2280

tgtccttgct ttggagactt taagacattt cctaaagcac aaataaaagc ctcgtatttc 2340

cccattgaga gttttgttcc aaggaatatg aagtgagaca tatgggtgag tcataataat 2400

caaaataatt tatgaagagc tgggtctgca atagctagtc taaaaactac ttgtgtgtca 2460

gtcctctggt tatagtatat aagagcctga ggaggtctgg caagatagat ggtgtattat 2520

ttatggatca ggctgctgca tacaaacctt gcatactatt atgcagctta cctaactctc 2580

agactattct gagtaatgct tgcttgctaa tgaatgtata ggagaccaca ttgtaattgt 2640

tcttagatga tggagtccat gcagtttctt agaaatcggt ctcagtgcat gctgtgcttt 2700

ttcacatttg ctctgggtta tctgggaagt atcaggttct gggaggcaac agcattaagt 2760

gataagaaaa ggagacattc tggcaaagcc aatctgctta aaggcaaagt ccagaacctg 2820

gaacctagag gcctttctct ctgcacgaaa aacaggtagt ttgcagtctg agatatggga 2880

gagcttttag gctacacagc aacccaaggg acctctcacc ttttgctgag cttcaatcag 2940

gaagctattt gcctggctcc agcagatgat gagataatga ggtagtgggt tttttattac 3000

tgttccattt tgcaacatcc tgcaacacca tcctgggaga caagagcatt acccagcttg 3060

gctttcacgg gggagggttg tattcagt 3088

<210> 41
<211> 3868
<212> DNA
<213> Homo Sapiens

<400> 41
atgaacctct gaaaactgcc ggcatctgag gtttcctcca aggccctctg aagtgcagcc 60

cataatgaag gtcttggcgg caggagttgt gcccctgctg ttggttctgc actggaaaca 120

tggggcgggg agccccctcc ccatcacccc tgtcaacgcc acctgtgcca tacgccaccc 180

atgtcacaac aacctcatga accagatcag gagccaactg gcacagctca atggcagtgc 240

caatgccctc tttattctct attacacagc ccagggggag ccgttcccca acaacctgga 300

caagctatgt ggccccaacg tgacggactt cccgcccttc cacgccaacg gcacggagaa 360

ggccaagctg gtggagctgt accgcatagt cgtgtacctt ggcacctccc tgggcaacat 420

cacccgggac cagaagatcc tcaaccccag tgccctcagc ctccacagca agctcaacgc 480

caccgccgac atcctgcgag gcctccttag caacgtgctg tgccgcctgt gcagcaagta 540

ccacgtgggc catgtggacg tgacctacgg ccctgacacc tcgggtaagg atgtcttcca 600

gaagaagaag ctgggctgtc aactcctggg gaagtataag cagatcatcg ccgtgttggc 660

ccaggccttc tagcaggagg tcttgaagtg tgctgtgaac cgagggatct caggagttgg 720

gtccagatgt gggggcctgt ccaagggtgg ctggggccca gggcatcgct aaacccaaat 780

gggggctgct ggcagacccc gagggtgcct ggccagtcca ctccactctg ggctgggctg 840

tgatgaagct gagcagagtg gaaacttcca tagggaggga gctagaagaa ggtgcccctt 900

cctctgggag attgtggact ggggagcgtg ggctggactt ctgcctctac ttgtcccttt 960

ggccccttgc tcactttgtg cagtgaacaa actacacaag tcatctacaa gagccctgac 1020

cacagggtga gacagcaggg cccaggggag tggaccagcc cccagcaaat tatcaccatc 1080

tgtgcctttg ctgcccctta ggttgggact taggtgggcc agaggggcta ggatcccaaa 1140

ggactccttg tcccctagaa gtttgatgag tggaagatag agaggggcct ctgggatgga 1200

aggctgtctt cttttgagga tgatcagaga acttgggcat aggaacaatc tggcagaagt 1260

ttccagaagg aggtcacttg gcattcaggc tcttggggag gcagagaagc caccttcagg 1320

cctgggaagg aagacactgg gaggaggaga ggcctggaaa gctttggtag gttcttcgtt 1380

ctcttccccg tgatcttccc tgcagcctgg gatggccagg gtctgatggc tggacctgca 1440

gcaggggttt gtggaggtgg gtagggcagg ggcaggttgc taagtcaggt gcagaggttc 1500

tgagggaccc aggctcttcc tctgggtaaa ggtctgtaag aaggggctgg ggtagctcag 1560

agtagcagct cacatctgag gccctgggag gtcttgtgag gtcacacaga ggtacttgag 1620

ggggactgga ggccgtctct ggtccccagg gcaagggaac agcagaactt agggtcaggg 1680

tctcagggaa ccctgagctc caagcgtgct gtgcgtctga cctggcatga tttctattta 1740

ttatgatatc ctatttatat taacttattg gtgctttcag tggccaagtt aattcccctt 1800

tccctggtcc ctactcaaca aaatatgatg atggctcccg acacaagcgc cagggccagg 1860

gcttagcagg gcctggtctg gaagtcgaca atgttacaag tggaataagc ttacgggtga 1920

agctcagaga agggtcggat ctgagagaat ggggaggcct gagtgggagt gggggggctt 1980

gctccacccc catcccctac tgtgacttgc tttagcgtgt cagggtccag gctgcagggg 2040

ctgggccaat ttgtggagag gccgggtgcc tttctgtctt gcttccaggg ggctggttca 2100

cactgttctt gggcgcccca gcattgtgtt gtgaggcgca ctgttcctgg cagatattgt 2160

gccccctgga gcagtgggca agacagtcct tgtggcccac cctgtccttg tttctgtgtc 2220

cccatgctgc ctctgaaata gcgccctgga acaaccctgc ccctgcaccc agcatgctcc 2280

gacacagcag ggaagctcct cctgtggccc ggacacccat agacggtgcg gggggcctgg 2340

ctgggccaga ccccaggaag gtggggtaga ctgggggggat cagctgccca ttgctcccaa 2400

gaggaggaga gggaggctgc agacgcctgg gactcagacc aggaagctgt gggccctcct 2460

gctccacccc catcccactc ccacccatgt ctgggctccc aggcagggaa cccgatctct 2520

tcctttgtgc tggggccagg cgagtggaga aacgccctcc agtctgagag caggggaggg 2580

aaggaggcag cagagttggg gcagctgctc agagcagtgt tctggcttct tctcaaaccc 2640

tgagcgggct gccggcctcc aagttcctcc gacaagatga tggtactaat tatggtactt 2700

ttcactcact ttgcaccttt ccctgtcgct ctctaagcac tttacctgga tggcgcgtgg 2760

gcagtgtgca ggcaggtcct gaggcctggg gttggggtgg agggtgcggc ccggagttgt 2820

ccatctgtcc atcccaacag caagacgagg atgtggctgt tgagatgtgg gccacactca 2880

cccttgtcca ggatgcaggg actgccttct ccttcctgct tcatccggct tagcttgggg 2940

ctggctgcat tcccccagga tgggcttcga gaaagacaaa cttgtctgga aaccagagtt 3000

gctgattcca cccgggggcc ccggctgact cgcccatcac ctcatctccc tgtggacttg 3060

ggagctctgt gccaggccca ccttgcggcc ctggctctga gtcgctctcc cacccagcct 3120

ggacttggcc ccatgggacc catcctcagt gctccctcca gatcccgtcc ggcagcttgg 3180

cgtccaccct gcacagcatc actgaatcac agagcctttg cgtgaaacag ctctgccagg 3240

ccgggagctg ggtttctctt ccctttttat ctgctggtgt ggaccacacc tgggcctggc 3300

cggaggaaga gagagtttac caagagagat gtctccgggc ccttatttat tatttaaaca 3360

ttttttaaa aagcactgct agtttacttg tctctcctcc ccatcgtccc catcgtcctc 3420

cttgtccctg acttggggca cttccaccct gacccagcca gtccagctct gccttgccgg 3480

ctctccagag tagacatagt gtgtggggtt ggagctctgg cacccgggga ggtagcattt 3540

ccctgcagat ggtacagatg ttcctgcctt agagtcatct ctagttcccc acctcaatcc 3600

cggcatccag ccttcagtcc cgcccacgtg ctagctccgt gggcccaccg tgcggcctta 3660

gaggtttccc tccttccttt ccactgaaaa gcacatggcc ttgggtgaca aattcctctt 3720

tgatgaatgt accctgtggg gatgtttcat actgacagat tatttttatt tattcaatgt 3780

catatttaaa atatttattt tttataccaa atgaatcact ttttttttta agaaaaaaaa 3840

gagaaatgaa taaagaatct actcttcg 3868

<210> 42
<211> 3145
<212> DNA
<213> Homo Sapiens

<400> 42
ccccaggtcc ggacaggccg agatgacgcc gagcccccctg ttgctgctcc tgctgccgcc 60

gctgctgctg ggggccttcc caccggccgc cgccgcccga ggcccccccaa agatggcgga 120

caaggtggtc ccacggcagg tggcccggct gggccgcact gtgcggctgc agtgcccagt 180

ggaggggac ccgccgccgc tgaccatgtg gaccaaggat ggccgcacca tccacagcgg 240
ctggagccgc ttccgcgtgc tgccgcaggg gctgaaggtg aagcaggtgg agcgggagga 300
tgccggcgtg tacgtgtgca aggccaccaa cggcttcggc agccttagcg tcaactacac 360
cctcgtcgtg ctggatgaca ttagcccagg gaaggagagc ctggggcccg acagctcctc 420
tgggggtcaa gaggaccccg ccagccagca gtgggcacga ccgcgcttca cacagccctc 480
caagatgagg cgccgggtga tcgcacggcc cgtgggtagc tccgtgcggc tcaagtgcgt 540
ggccagcggg caccctcggc ccgacatcac gtggatgaag gacgaccagg ccttgacgcg 600
cccagaggcc gctgagccca ggaagaagaa gtggacactg agcctgaaga acctgcggcc 660
ggaggacagc ggcaaataca cctgccgcgt gtcgaaccgc gcgggcgcca tcaacgccac 720
ctacaaggtg gatgtgatcc agcggacccg ttccaagccc gtgctcacag gcacgcaccc 780
cgtgaacacg acggtggact tcgggggggac cacgtccttc cagtgcaagg tgcgcagcga 840
cgtgaagccg gtgatccagt ggctgaagcg cgtggagtac ggcgccgagg gccgccacaa 900
ctccaccatc gatgtgggcg gccagaagtt tgtggtgctg cccacgggtg acgtgtggtc 960
gcggcccgac ggctcctacc tcaataagct gctcatcacc cgtgcccgcc aggacgatgc 1020
gggcatgtac atctgccttg gcgccaacac catgggctac agcttccgca gcgccttcct 1080
caccgtgctg ccagacccaa aaccgcaagg gccacctgtg gcctcctcgt cctcggccac 1140
tagcctgccg tggcccgtgg tcatcggcat cccagccggc gctgtcttca tcctgggcac 1200
cctgctcctg tggctttgcc aggcccagaa gaagccgtgc acccccgcgc ctgcccctcc 1260
cctgcctggg caccgcccgc cggggacggc ccgcgaccgc agcggagaca aggaccttcc 1320
ctcgttggcc gccctcagcg ctggccctgg tgtggggctg tgtgaggagc atgggtctcc 1380
ggcagccccc cagcacttac tgggcccagg cccagttgct ggccctaagt tgtaccccaa 1440
actctacaca gacatccaca cacacacaca cacacactct cacacacact cacacgtgga 1500
gggcaaggtc caccagcaca tccactatca gtgctagacg gcaccgtatc tgcagtgggc 1560
acgggggggc cggccagaca ggcagactgg gaggatggag gacggagctg cagacgaagg 1620
caggggaccc atggcgagga ggaatggcca gcaccccagg cagtctgtgt gtgaggcata 1680
gcccctggac acacacacac agacacacac actacctgga tgcatgtatg cacacacatg 1740
cgcgcacacg tgctccctga aggcacacgt acgcacacac gcacatgcac agatatgccg 1800
cctgggcaca cagataagct gcccaaatgc acgcacacgc acagagacat gccagaacat 1860
acaaggacat gctgcctgaa catacacacg cacacccatg cgcagatgtg ctgcctggac 1920
acacacacac acacggatat gctgtctgga cgcacacacg tgcagatatg gtatccggac 1980

acacacgtgc acagatatgc tgcctggaca cacagataat gctgccttga cacacacatg 2040

cacggatatt gcctggacac acacacacac acgcgtgcac agatatgctg tctggacagg 2100

cacacacatg cagatatgct gcctggacac acacttccag acacacgtgc acaggcgcag 2160

atatgctgcc tggacacacg cagatatgct gtctagtcac acacacacgc agacatgctg 2220

tccggacaca cacacgcatg cacagatatg ctgtccggac acacacacgc acgcagatat 2280

gctgcctgga cacacacaca gataatgctg cctcaacact cacacacgtg cagatattgc 2340

ctggacacac acatgtgcac agatatgctg tctggacatg cacacacgtg cagatatgct 2400

gtccggatac acacgcacgc acacatgcag atatgctgcc tgggcacaca cttccggaca 2460

cacatgcaca cacaggtgca gatatgctgc ctggacacac gcagactgac gtgcttttgg 2520

gagggtgtgc cgtgaagcct gcagtacgtg tgccgtgagg ctcatagttg atgagggact 2580

ttccctgctc caccgtcact cccccaactc tgcccgcctc tgtccccgcc tcagtccccg 2640

cctccatccc cgcctctgtc ccctggcctt ggcggctatt tttgccacct gccttgggtg 2700

cccaggagtc ccctactgct gtgggctggg gttgggggca cagcagcccc aagcctgaga 2760

ggctggagcc catggctagt ggctcatccc cactgcattc tccccctgac acagagaagg 2820

ggccttggta tttatattta agaaatgaag ataatattaa taatgatgga aggaagactg 2880

ggttgcaggg actgtggtct ctcctggggc ccgggacccg cctggtcttt cagccatgct 2940

gatgaccaca ccccgtccag gccagacacc acccccacc ccactgtcgt ggtggcccca 3000

gatctctgta attttatgta gagtttgagc tgaagccccg tatatttaat ttattttgtt 3060

aaacatgaaa gtgcatcctt tccctccaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 3120

aaaaaaaaaa aaaaaaaaaa aaaaa 3145

<210> 43
<211> 3273
<212> DNA
<213> Homo Sapiens

<400> 43
tgaaaggcgg ttgtggtgca aaggaaaacc cacaggccaa ggaatgggaa gaccaaggtt 60

gacacttgtt tgtcacgtgt caataatcat ctctgcccgg gacctcagca tgaacaacct 120

cacagagctt cagcctggcc tcttccacca cctgcgcttc ttggaggagc tgcgtctctc 180

tgggaaccat ctctcacaca tcccaggaca agcattctct ggtctctaca gcctgaaaat 240

cctgatgctg cagaacaatc agctgggagg aatccccgca gaggcgctgt gggagctgcc 300

gagcctgcag tcgctgcgcc tagatgccaa cctcatctcc ctggtcccgg agaggagctt 360

tgaggggctg tcctccctcc gccacctctg gctggacgac aatgcactca cggagatccc 420

tgtcagggcc ctcaacaacc tccctgccct gcaggccatg accctggccc tcaaccgcat 480
cagccacatc cccgactacg cgttccagaa tctcaccagc cttgtggtgc tgcatttgca 540
taacaaccgc atccagcatc tggggaccca cagcttcgag gggctgcaca atctggagac 600
actagacctg aattataaca agctgcagga gttccctgtg gccatccgga ccctgggcag 660
actgcaggaa ctggggttcc ataacaacaa catcaaggcc atcccagaaa aggccttcat 720
ggggaaccct ctgctacaga cgatacactt ttatgataac ccaatccagt ttgtgggaag 780
atcggcattc cagtacctgc ctaaactcca cacactatct ctgaatggtg ccatggacat 840
ccaggagttt ccagatctca aaggcaccac cagcctggag atcctgaccc tgacccgcgc 900
aggcatccgg ctgctcccat cggggatgtg ccaacagctg cccaggctcc gagtcctgga 960
actgtctcac aatcaaattg aggagctgcc cagcctgcac aggtgtcaga aattggagga 1020
aatcggcctc caacacaacc gcatctggga aattggagct gacaccttca gccagctgag 1080
ctccctgcaa gccctggatc ttagctggaa cgccatccgg tccatccacc ccgaggcctt 1140
ctccaccctg cactccctgg tcaagctgga cctgacagac aaccagctga ccacactgcc 1200
cctggctgga cttgggggct tgatgcatct gaagctcaaa gggaaccttg ctctctccca 1260
ggccttctcc aaggacagtt tcccaaaact gaggatcctg gaggtgcctt atgcctacca 1320
gtgctgtccc tatgggatgt gtgccagctt cttcaaggcc tctgggcagt gggaggctga 1380
agaccttcac cttgatgatg aggagtcttc aaaaaggccc ctgggcctcc ttgccagaca 1440
agcagagaac cactatgacc aggacctgga tgagctccag ctggagatgg aggactcaaa 1500
gccacacccc agtgtccagt gtagccctac tccaggcccc ttcaagccct gtgagtacct 1560
ctttgaaagc tggggcatcc gcctggccgt gtgggccatc gtgttgctct ccgtgctctg 1620
caatggactg gtgctgctga ccgtgttcgc tggcgggcct gtcccctgc ccccggtcaa 1680
gtttgtggta ggtgcgattg caggcgccaa caccttgact ggcatttcct gtggccttct 1740
agcctcagtc gatgccctga cctttggtca gttctctgag tacggagccc gctgggagac 1800
ggggctaggc tgccgggcca ctggcttcct ggcagtactt gggtcggagg catcggtgct 1860
gctgctcact ctggccgcag tgcagtgcag cgtctccgtc tcctgtgtcc gggcctatgg 1920
gaagtccccc tccctgggca gcgttcgagc aggggtccta ggctgcctgg cactggcagg 1980
gctggccgcc gcgctgcccc tggcctcagt gggagaatac ggggcctccc cactctgcct 2040
gccctacgcg ccacctgagg gtcagccagc agccctgggc ttcaccgtgg ccctggtgat 2100
gatgaactcc ttctgtttcc tggtcgtggc cggtgcctac atcaaactgt actgtgacct 2160
gccgcggggc gactttgagg ccgtgtggga ctgcgccatg gtgaggcacg tggcctggct 2220
catcttcgca gacgggctcc tctactgtcc cgtggccttc ctcagctttg cctccatgct 2280

gggcctcttc cctgtcacgc ccgaggccgt caagtctgtc ctgctggtgg tgctgcccct 2340

gcctgcctgc ctcaacccac tgctgtacct gctcttcaac ccccacttcc gggatgacct 2400

tcggcggctt cggccccgcg caggggactc agggccccta gcctatgctg cggccgggga 2460

gctggagaag agctcctgtg attctaccca ggccctggta gccttctctg atgtggatct 2520

cattctggaa gcttctgaag ctgggcggcc ccctgggctg gagacctatg gcttcccctc 2580

agtgaccctc atctcctgtc agcagccagg ggcccccagg ctggagggca gccattgtgt 2640

agagccagag gggaaccact ttgggaaccc ccaaccctcc atggatggag aactgctgct 2700

gagggcagag ggatctacgc cagcaggtgg aggcttgtca gggggtggcg gctttcagcc 2760

ctctggcttg gcctttgctt cacacgtgta aatatccctc cccattcttc tcttccctc 2820

tcttcccttt cctctctccc cctcggtgaa tgatggctgc ttctaaaaca aatacaacca 2880

aaactcagca gtgtgatcta tagcaggatg gcccagtccc tggctccact gatcacctct 2940

ctcctgtgac catcaccaac gggtgcctct tggcctggct ttcccttggc cttcctcagc 3000

ttcaccttga tactgggcct cttccttgtc atgtctgaag ctgtggacca gagacctgga 3060

cttttgtctg cttaagggaa atgagggaag taaagacagt gaaggggtgg agggttgatc 3120

agggcacagt ggacagggag acctcacaga gaaaggcctg gaaggtgatt tcccgtgtga 3180

ctcatggata ggatacaaaa tgtgttccat gtaccattaa tcttgacata tgccatgcat 3240

aaagacttcc tattaaaata agctttggaa gag 3273

<210> 44
<211> 2192
<212> DNA
<213> Homo Sapiens

<400> 44
agtctggccc tggacaaccc cagcaaagcc gccctcagcc agcccagaag cactgggcct 60

tggccacagc aacacccact gagcacgctg ggagctgagt atggcgtccc tggtctcgct 120

ggagctgggg ctgcttctgg ctgtgctggt ggtgacggcg acggcgtccc cgcctgctgg 180

tctgctgagc ctgctcacct ctggccaggg cgctctggat caagaggctc tgggcggcct 240

gttaaatacg ctggcggacc gtgtgcactg caccaacggg ccgtgtggaa agtgcctgtc 300

tgtggaggac gccctgggcc tgggcgagcc tgaggggtca gggctgcccc cgggcccggt 360

cctggaggcc aggtacgtcg cccgcctcag tgccgccgcc gtcctgtacc tcagcaaccc 420

cgagggcacc tgtgaggaca ctcgggctgg cctctgggcc tctcatgcag accacctcct 480

ggccctgctc gagagcccca aggccctgac cccgggcctg agctggctgc tgcagaggat 540

gcaggcccgg gctgccggcc agacccccaa gacggcctgc gtagatatcc ctcagctgct 600

ggaggaggcg gtgggggcgg gggctccggg cagtgctggc ggcgtcctgg ctgccctgct 660
ggaccatgtc aggagcgggt cttgcttcca cgccttgccg agccctcagt acttcgtgga 720
ctttgtgttc cagcagcaca gcagcgaggt ccctatgacg ctggccgagc tgtcagcctt 780
gatgcagcgc ctgggggtgg gcagggaggc ccacagtgac cacagtcatc ggcacagggg 840
agccagcagc cgggaccctg tgcccctcat cagctccagc aacagctcca gtgtgtggga 900
cacggtatgc ctgagtgcca gggacgtgat ggctgcatat ggactgtcgg aacaggctgg 960
ggtgaccccg gaggcctggg cccaactgag ccctgccctg ctccaacagc agctgagtgg 1020
agcctgcacc tcccagtcca ggcccccccgt ccaggaccag ctcagccagt cagagaggta 1080
tctgtacggc tccctggcca cgctgctcat ctgcctctgc gcggtctttg gcctcctgct 1140
gctgacctgc actggctgca ggggggtcgc ccactacatc ctgcagacct tcctgagcct 1200
ggcagtgggt gcactcactg gggacgctgt cctgcatctg acgcccaagg tgctggggct 1260
gcatacacac agcgaagagg gcctcagccc acagcccacc tggcgcctcc tggctatgct 1320
ggccgggctc tacgccttct tcctgtttga gaacctcttc aatctcctgc tgcccaggga 1380
cccggaggac ctggaggacg ggccctgcgg ccacagcagc catagccacg ggggccacag 1440
ccacggtgtg tccctgcagc tggcacccag cgagctccgg cagcccaagc ccccccacga 1500
gggctcccgc gcagacctgg tggcggagga gagcccggag ctgctgaacc ctgagcccag 1560
gagactgagc ccagagttga ggctactgcc ctatatgatc actctgggcg acgccgtgca 1620
caacttcgcc gacgggctgg ccgtgggcgc cgccttcgcg tcctcctgga agaccgggct 1680
ggccacctcg ctggccgtgt tctgccacga gttgccacac gagctggggg acttcgccgc 1740
cttgctgcac gcggggctgt ccgtgcgcca agcactgctg ctgaacctgg cctccgcgct 1800
cacggccttc gctggtctct acgtggcact cgcggttgga gtcagcgagg agagcgaggc 1860
ctggatcctg gcagtggcca ccggcctgtt cctctacgta gcactctgcg acatgctccc 1920
ggcgatgttg aaagtacggg acccgcggcc ctggctcctc ttcctgctgc acaacgtggg 1980
cctgctgggc ggctggaccg tcctgctgct gctgtccctg tacgaggatg acatcacctt 2040
ctgataccct gccctagtcc cccacctttg acttaagatc ccacacctca caaacctaca 2100
gcccagaaac cagaagcccc tatagaggcc ccagtcccaa ctccagtaaa gacactcttg 2160
tccttggaaa aaaaaaaaaa aaaaaaaaaa aa 2192

<210> 45
<211> 3014
<212> DNA
<213> Homo Sapiens

<400> 45
aagctcgggc tccggcacgt agttgggaaa cttgcgggtc ctagaagtcg cctccccgcc 60
ttgccggccg cccttgcagc cccgagccga gcagcaaagt gagacattgt gcgcctgcca 120
gatccgccgg ccgcggaccg gggctgcctc ggaaacacag aggggtcttc tctcgccctg 180
catataatta gcctgcacac aaagggagca gctgaatgga ggttgtcact ctctggaaaa 240
ggatttctga ccgagcgctt ccaatggaca ttctccagtc tctctggaaa gattctcgct 300
aatggatttc ctgctgctcg gtctctgtct atactggctg ctgaggaggc cctcgggggt 360
ggtcttgtgt ctgctggggg cctgctttca gatgctgccc gccgcccccа gcgggtgccc 420
gcagctgtgc cggtgcgagg ggcggctgct gtactgcgag gcgctcaacc tcaccgaggc 480
gccccacaac ctgtccggcc tgctgggctt gtccctgcgc tacaacagcc tctcggagct 540
gcgcgccggc cagttcacgg ggttaatgca gctcacgtgg ctctatctgg atcacaatca 600
catctgctcc gtgcaggggg acgcctttca gaaactgcgc cgagttaagg aactcacgct 660
gagttccaac cagatcaccc aactgcccaa caccaccttc cggcccatgc ccaacctgcg 720
cagcgtggac ctctcgtaca acaagctgca ggcgctcgcg cccgacctct tccacgggct 780
gcggaagctc accacgctgc atatgcgggc caacgccatc cagtttgtgc ccgtgcgcat 840
cttccaggac tgccgcagcc tcaagtttct cgacatcgga tacaatcagc tcaagagtct 900
ggcgcgcaac tctttcgccg gcttgtttaa gctcaccgag ctgcacctcg agcacaacga 960
cttggtcaag gtgaacttcg cccacttccc gcgcctcatc tccctgcact cgctctgcct 1020
gcggaggaac aaggtggcca ttgtggtcag ctcgctggac tgggtttgga acctggagaa 1080
aatggacttg tcgggcaacg agatcgagta catggagccc catgtgttcg agaccgtgcc 1140
gcacctgcag tccctgcagc tggactccaa ccgcctcacc tacatcgagc cccggatcct 1200
caactcttgg aagtccctga caagcatcac cctggccggg aacctgtggg attgcgggcg 1260
caacgtgtgt gccctagcct cgtggctcaa caacttccag gggcgctacg atggcaactt 1320
gcagtgcgcc agcccggagt acgcacaggg cgaggacgtc ctggacgccg tgtacgcctt 1380
ccacctgtgc gaggatgggg ccgagcccac cagcggccac ctgctctcgg ccgtcaccaa 1440
ccgcagtgat ctggggcccc ctgccagctc ggccaccacg ctcgcggacg gcggggaggg 1500
gcagcacgac ggcacattcg agcctgccac cgtggctctt ccaggcggcg agcacgccga 1560
gaacgccgtg cagatccaca aggtggtcac gggcaccatg gccctcatct tctccttcct 1620
catcgtggtc ctggtgctct acgtgtcctg gaagtgtttc ccagccagcc tcaggcagct 1680
cagacagtgc tttgtcacgc agcgcaggaa gcaaaagcag aaacagacca tgcatcagat 1740
ggctgccatg tctgcccagg aatactacgt tgattacaaa ccgaaccaca ttgagggagc 1800

cctggtgacc atcaacgagt atggctcgtg tacctgccac cagcagcccg cgagggaatg 1860

cgaggtgtga ttgtcccagt ggctctcaac ccatgcgcta ccaaatacgc ctgggcagcc 1920

gggacgggcc ggcgggcacc aggctggggt ctccttgtct gtgctctgat atgctccttg 1980

actgaaactt taaggggatc tctcccagag acttgacatt ttagctttat tgtgtcttaa 2040

aaacaaaagc gaattaaaac acaacaaaaa accccacccc acaaccttca ggacagtcta 2100

tcttaaattt catatgagaa ctccttcctc cctttgaaga tctgtccata ttcaggaatc 2160

tgagagtgta aaaaagttct gtcaacagat tcagtcccca gtcctccttc tctgagcaga 2220

cctctcatcc gcagcatata tctgctagat cttctaagag ctgaaatgga gacttctgga 2280

agaaggtggg aaagaaatcc atctcggctt aattaaccat ttatcgcatc atattactcc 2340

catcttaaaa gtgcacgcgt tgtttttctg aaccctcaca caaaggctac tactgtggtc 2400

ccatatctgt cggcccatga gaaacagtgt tcttggacct cacagccaag cagcactgaa 2460

ctgcagcaaa atccagcaac acattcagca gcgagcagcc tgctgagctc cactggttta 2520

tccggggcca ccaaccccaa agaactggga tgaaagcaga tgtgagagag gaaaaggatc 2580

tgtttttgtt ttattttcta ccaggcccag ctctttgtgg ggggaataaa aaagaagaaa 2640

aatcgagctc caagctggtg ccctgccaag cttcctcccc tcccttccta gtccaagcac 2700

tccaccgtct gtgcagactg cataacagca atttctggaa acaggctgaa gaatctgggc 2760

cagtccagag gcagtggatt cctggtttat gtgtggtggg gtttttagga attttatttt 2820

tcaccttaat tctttcaaca actgccagct gtttgaagca catctgtaat aaacagcttc 2880

tgtttgtaaa atgagactga agttatcctc tccagagaaa ttcctgaatc ttctctgtag 2940

ttcaatgcct tcactgacag tttggctcaa aaagtatgag tgtggtaaat attaaagaat 3000

gttaatacaa gtgt 3014

<210> 46
<211> 1128
<212> DNA
<213> Homo Sapiens

<400> 46
atggcgaacg cgagcgagcc gggtggcagc ggcggcggcg aggcggccgc cctgggcctc 60

aagctggcca cgctcagcct gctgctgtgc gtgagcctag cgggcaacgt gctgttcgcg 120

ctgctgatcg tgcgggagcg cagcctgcac cgcgccccgt actacctgct gctcgacctg 180

tgcctggccg acgggctgcg cgcgctcgcc tgcctcccgg ccgtcatgct ggcggcgcgg 240

cgtgcggcgg ccgcggcggg ggcgccgccg ggcgcgctgg gctgcaagct gctcgccttc 300

ctggccgcgc tcttctgctt ccacgccgcc ttcctgctgc tgggcgtggg cgtcacccgc 360

tacctggcca tcgcgcacca ccgcttctat gcagagcgcc tggccggctg gccgtgcgcc 420

gccatgctgg tgtgcgccgc ctgggcgctg gcgctggccg cggccttccc gccagtgctg 480

gacggcggtg gcgacgacga ggacgcgccg tgcgccctgg agcagcggcc cgacggcgcc 540

cccggcgcgc tgggcttcct gctgctgctg gccgtggtgg tgggcgccac gcacctcgtc 600

tacctccgcc tgctcttctt catccacgac cgccgcaaga tgcggcccgc gcgcctggtg 660

cccgccgtca gccacgactg gaccttccac ggcccgggcg ccaccggcca ggcggccgcc 720

aactggacgg cgggcttcgg ccgcgggccc acgccgcccg cgcttgtggg catccggccc 780

gcagggccgg gccgcggcgc gcgccgcctc ctcgtgctgg aagaattcaa gacggagaag 840

aggctgtgca agatgttcta cgccgtcacg ctgctcttcc tgctcctctg ggggccctac 900

gtcgtggcca gctacctgcg ggtcctggtg cggcccggcg ccgtcccccа ggcctacctg 960

acggcctccg tgtggctgac cttcgcgcag gccggcatca accccgtcgt gtgcttcctc 1020

ttcaacaggg agctgaggga ctgcttcagg gcccagttcc cctgctgcca gagccccccgg 1080

accacccagg cgacccatcc ctgcgacctg aaaggcattg gtttatga 1128

<210> 47
<211> 1736
<212> DNA
<213> Homo Sapiens

<400> 47
gagggagggg cgggggctgg aggcagcagc gcccccgcac tccccgcgtc tcgcacactt 60

gcaccggtcg ctcgcgcgca gcccggcgtc gccccacgcc gcgctcgctc ctccctccct 120

cctcccgctc cgtggctccc gtgctcctgg cgaggctcag gcgcggagcg cgcggacggg 180

cgcaccgaca gacggccccg gggacgcctc ggctcgcgcc tcccgggcgg gctatgttga 240

ttgccccgcc ggggccggcc cgcgggatca gcacagcccg gcccgcggcc ccggcggcca 300

atcgggacta tgaaccggaa agcgcggcgc tgcctgggcc acctctttct cagcctgggc 360

atggtctacc tccggatcgg tggcttctcc tcagtggtag ctctgggcgc aagcatcatc 420

tgtaacaaga tcccaggcct ggctcccaga cagcgggcga tctgccagag ccggcccgac 480

gccatcatcg tcataggaga aggctcacaa atgggcctgg acgagtgtca gtttcagttc 540

cgcaatggcc gctggaactg ctctgcactg ggagagcgca ccgtcttcgg gaaggagctc 600

aaagtgggga gccgggaggc tgcgttcacc tacgccatca ttgccgccgg cgtggcccac 660

gccatcacag ctgcctgtac ccagggcaac ctgagcgact gtggctgcga caaagagaag 720

caaggccagt accaccggga cgagggctgg aagtggggtg gctgctctgc cgacatccgc 780

tacggcatcg gcttcgccaa ggtctttgtg gatgcccggg agatcaagca gaatgcccgg 840

actctcatga acttgcacaa caacgaggca ggccgaaaga tcctggagga gaacatgaag 900

ctggaatgta agtgccacgg cgtgtcaggc tcgtgcacca ccaagacgtg ctggaccaca 960

ctgccacagt ttcgggagct gggctacgtg ctcaaggaca agtacaacga ggccgttcac 1020

gtggagcctg tgcgtgccag ccgcaacaag cggcccacct tcctgaagat caagaagcca 1080

ctgtcgtacc gcaagcccat ggacacggac ctggtgtaca tcgagaagtc gcccaactac 1140

tgcgaggagg acccggtgac cggcagtgtg ggcacccagg gccgcgcctg caacaagacg 1200

gctccccagg ccagcggctg tgacctcatg tgctgtgggc gtggctacaa cacccaccag 1260

tacgcccgcg tgtggcagtg caactgtaag ttccactggt gctgctatgt caagtgcaac 1320

acgtgcagcg agcgcacgga gatgtacacg tgcaagtgag ccccgtgtgc acaccaccct 1380

cccgctgcaa gtcagattgc tgggaggact ggaccgtttc caagctgcgg gctccctggc 1440

aggatgctga gcttgtcttt tctgctgagg agggtacttt tcctgggttt cctgcaggca 1500

tccgtggggg aaaaaaaatc tctcagagcc ctcaactatt ctgttccaca cccaatgctg 1560

ctccaccctc ccccagacac agcccaggtc cctccgcggc tggagcgaag ccttctgcag 1620

caggaactct ggacccctgg gcctcatcac agcaatattt aacaatttat tctgataaaa 1680

ataatattaa tttatttaat taaaaagaat tcttccacaa aaaaaaaaaa aaaaaa 1736

<210> 48
<211> 3195
<212> DNA
<213> Homo Sapiens

<400> 48
acagcatgga gtggggttac ctgttggaag tgacctcgct gctggccgcc ttggcgctgc 60

tgcagcgctc tagcggcgct gcggccgcct cggccaagga gctggcatgc caagagatca 120

ccgtgccgct gtgtaagggc atcggctaca actacaccta catgcccaat cagttcaacc 180

acgacacgca agacgaggcg ggcctggagg tgcaccagtt ctggccgctg gtggagatcc 240

agtgctcgcc cgatctcaag ttcttcctgt gcagcatgta cacgcccatc tgcctagagg 300

actacaagaa gccgctgccg ccctgccgct cggtgtgcga gcgcgccaag gccggctgcg 360

cgccgctcat gcgccagtac ggcttcgcct ggcccgaccg catgcgctgc gaccggctgc 420

ccgagcaagg caaccctgac acgctgtgca tggactacaa ccgcaccgac ctaaccaccg 480

ccgcgcccag cccgccgcgc cgcctgccgc cgccgccgcc cggcgagcag ccgccttcgg 540

gcagcggcca cggccgcccg ccgggggcca ggcccccgca ccgcggaggc ggcaggggcg 600

gtggcggcgg ggacgcggcg gcgcccccag ctcgcggcgg cggcggtggc gggaaggcgc 660

ggcccccctgg cggcggcgcg gctccctgcg agcccgggtg ccagtgccgc gcgcctatgg 720

tgagcgtgtc cagcgagcgc cacccgctct acaaccgcgt caagacaggc cagatcgcta 780

actgcgcgct gccctgccac aacccctttt tcagccagga cgagcgcgcc ttcaccgtct 840

tctggatcgg cctgtggtcg gtgctctgct tcgtgtccac cttcgccacc gtctccacct 900

tccttatcga catggagcgc ttcaagtacc cggagcggcc cattatcttc ctctcggcct 960

gctacctctt cgtgtcggtg ggctacctag tgcgcctggt ggcgggccac gagaaggtgg 1020

cgtgcagcgg tggcgcgccg ggcgcggggg gcgctggggg cgcgggcggc gcggcggcgg 1080

gcgcgggcgc ggcgggcgcg ggcgcgggcg gcccgggcgg gcgcggcgag tacgaggagc 1140

tgggcgcggt ggagcagcac gtgcgctacg agaccaccgg ccccgcgctg tgcaccgtgg 1200

tcttcttgct ggtctacttc ttcggcatgg ccagctccat ctggtgggtg atcttgtcgc 1260

tcacatggtt cctggcggcc ggtatgaagt ggggcaacga agccatcgcc ggctactcgc 1320

agtacttcca cctggccgcg tggcttgtgc ccagcgtcaa gtccatcgcg gtgctggcgc 1380

tcagctcggt ggacggcgac ccggtggcgg gcatctgcta cgtgggcaac cagagcctgg 1440

acaacctgcg cggcttcgtg ctggcgccgc tggtcatcta cctcttcatc ggcaccatgt 1500

tcctgctggc cggcttcgtg tccctgttcc gcatccgctc ggtcatcaag caacaggacg 1560

gccccaccaa gacgcacaag ctggagaagc tgatgatccg cctgggcctg ttcaccgtgc 1620

tctacaccgt gcccgccgcg gtggtggtcg cctgcctctt ctacgagcag cacaaccgcc 1680

cgcgctggga ggccacgcac aactgcccgt gcctgcggga cctgcagccc gaccaggcac 1740

gcaggcccga ctacgccgtc ttcatgctca agtacttcat gtgcctagtg gtgggcatca 1800

cctcgggcgt gtgggtctgg tccggcaaga cgctggagtc ctggcgctcc ctgtgcaccc 1860

gctgctgctg ggccagcaag ggcgccgcgg tgggcggggg cgcgggcgcc acggccgcgg 1920

ggggtggcgg cgggccgggg ggcggcggcg gcgggggacc cggcggcggc ggggggccgg 1980

gcggcggcgg gggctccctc tacagcgacg tcagcactgg cctgacgtgg cggtcgggca 2040

cggcgagctc cgtgtcttat ccaaagcaga tgccattgtc ccaggtctga gcggagggga 2100

gggggcgccc aggaggggtg gggagggggg cgaggagacc caagtgcagc gaagggacac 2160

ttgatgggct gaggttccca ccccttcaca gtgttgattg ctattagcat gataatgaac 2220

tcttaatggt atccattagc tgggacttaa atgactcact tagaacaaag tacctggcat 2280

tgaagcctcc cagacccagc ccctttttcct ccattgatgt gcggggagct cctcccgcca 2340

cgcgttaatt tctgttggct gaggagggtg gactctgcgg cgtttccaga acccgagatt 2400

tggagccctc cctggctgca cttggctggg tttgcagtca gatacacaga tttcacctgg 2460

gagaacctct ttttctccct cgactcttcc tacgtaaact cccacccctg acttaccctg 2520

gaggaggggt gaccgccacc tgatgggatt gcacggtttg ggtattctta atgaccaggc 2580

aaatgcctta agtaaacaaa caagaaatgt cttaattata caccccacgt aaatacgggt 2640

ttcttacatt agaggatgta tttatataat tatttgttaa attgtaaaaa aaaaaagtgt 2700

aaaatatgta tatatccaaa gatatagtgt gtacattttt ttgtaaaaag tttagaggct 2760

taccccctgta agaacagata taagtattct attttgtcaa taaaatgact tttgataaat 2820

gatttaacca ttgccctctc ccccgcctct tctgagctgt caccttaaa gtgcttgcta 2880

aggacgcatg gggaaaatgg acattttctg gcttgtcatt ctgtacactg accttaggca 2940

tggagaaaat tacttgttaa actctagttc ttaagttgtt agccaagtaa atatcattgt 3000

tgaactgaaa tcaaaattga gtttttgcac cttccccaaa gacggtgttt ttcatgggag 3060

ctcttttctg atccatggat aacaactctc actttagtgg atgtaaatgg aacttctgca 3120

aggcagtaat tccccttagg ccttgttatt tatcctgcat ggtatcacta aaggtttcaa 3180

aaccctgaaa aaaaa 3195

<210> 49
<211> 1380
<212> DNA
<213> Homo Sapiens

<400> 49
ccgggtcgga gccccccgga gctgcgcgcg ggcttgcagc gcctcgcccg cgctgtcctc 60

ccggtgtccc gcttctccgc gccccagccg ccggctgcca gcttttcggg gccccgagtc 120

gcacccagcg aagagagcgg gcccgggaca agctcgaact ccggccgcct cgcccttccc 180

cggctccgct ccctctgccc cctcggggtc gcgcgcccac gatgctgcag ggccctggct 240

cgctgctgct gctcttcctc gcctcgcact gctgcctggg ctcggcgcgc gggctcttcc 300

tctttggcca gcccgacttc tcctacaagc gcagcaattg caagcccatc cctgccaacc 360

tgcagctgtg ccacggcatc gaataccaga acatgcggct gcccaacctg ctgggccacg 420

agaccatgaa ggaggtgctg gagcaggccg gcgcttggat cccgctggtc atgaagcagt 480

gccacccgga caccaagaag ttcctgtgct cgctcttcgc ccccgtctgc ctcgatgacc 540

tagacgagac catccagcca tgccactcgc tctgcgtgca ggtgaaggac cgctgcgccc 600

cggtcatgtc cgccttcggc ttcccctggc ccgacatgct tgagtgcgac cgtttccccc 660

aggacaacga cctttgcatc cccctcgcta gcagcgacca cctcctgcca gccaccgagg 720

aagctccaaa ggtatgtgaa gcctgcaaaa ataaaaatga tgatgacaac gacataatgg 780

aaacgctttg taaaaatgat tttgcactga aaataaaagt gaaggagata acctacatca 840

accgagatac caaaatcatc ctggagacca agagcaagac catttacaag ctgaacggtg 900

tgtccgaaag ggacctgaag aaatcggtgc tgtggctcaa agacagcttg cagtgcacct 960

gtgaggagat gaacgacatc aacgcgccct atctggtcat gggacagaaa cagggtgggg 1020

agctggtgat cacctcggtg aagcggtggc agaaggggca gagagagttc aagcgcatct 1080

cccgcagcat ccgcaagctg cagtgctagt cccggcatcc tgatggctcc gacaggcctg 1140

ctccagagca cggctgacca tttctgctcc gggatctcag ctcccgttcc ccaagcacac 1200

tcctagctgc tccagtctca gcctgggcag cttccccctg ccttttgcac gtttgcatcc 1260

ccagcatttc ctgagttata aggccacagg agtggatagc tgttttcacc taaaggaaaa 1320

gcccacccga atcttgtaga aatattcaaa ctaataaaat catgaatatt tttatgaagt 1380

<210> 50
<211> 2573
<212> DNA
<213> Homo Sapiens

<400> 50
gaggagggac ctacaaagac tggaaactat tcttagctcc gtcactgact ccaagttcat 60

cccctctgtc tttcagtttg gttgagatat aggctactct tcccaactca gtcttgaaga 120

gtatcaccaa ctgcctcatg tgtggtgacc ttcactgtcg tatgccagtg actcatctgg 180

agtaatctca acaacgagtt accaatactt gctcttgatt gataaacaga atggggtttt 240

ggatcttagc aattctcaca attctcatgt attccacagc agcaaagttt agtaaacaat 300

catggggcct ggaaaatgag gctttaattg taagatgtcc tagacaagga aaacctagtt 360

acaccgtgga ttggtattac tcacaaacaa acaaaagtat tcccactcag gaaagaaatc 420

gtgtgtttgc ctcaggccaa cttctgaagt ttctaccagc tgcagttgct gattctggta 480

tttatacctg tattgtcaga agtcccacat tcaataggac tggatatgcg aatgtcacca 540

tatataaaaa acaatcagat tgcaatgttc cagattattt gatgtattca acagtatctg 600

gatcagaaaa aaattccaaa atttattgtc ctaccattga cctctacaac tggacagcac 660

ctcttgagtg gtttaagaat tgtcaggctc ttcaaggatc aaggtacagg gcgcacaagt 720

catttttggt cattgataat gtgatgactg aggacgcagg tgattacacc tgtaaattta 780

tacacaatga aaatggagcc aattatagtg tgacggcgac caggtccttc acggtcaagg 840

atgagcaagg cttttctctg tttccagtaa tcggagcccc tgcacaaaat gaaataaagg 900

aagtggaaat tggaaaaaac gcaaacctaa cttgctctgc ttgttttgga aaaggcactc 960

agttcttggc tgccgtcctg tggcagctta atggaacaaa aattacagac tttggtgaac 1020

caagaattca acaagaggaa gggcaaaatc aaagtttcag caatgggctg gcttgtctag 1080

acatggtttt aagaatagct gacgtgaagg aagaggattt attgctgcag tacgactgtc 1140

tggccctgaa tttgcatggc ttgagaaggc acaccgtaag actaagtagg aaaaatccaa 1200

gtaaggagtg tttctgagac tttgatcacc tgaactttct ctagcaagtg taagcagaat 1260

ggagtgtggt tccaagagat ccatcaagac aatgggaatg gcctgtgcca taaaatgtgc 1320

ttctcttctt cgggatgttg tttgctgtct gatctttgta gactgttcct gtttgctggg 1380

agcttctctg ctgcttaaat tgttcgtcct cccccactcc ctcctatcgt tggtttgtct 1440

agaacactca gctgcttctt tggtcatcct tgttttctaa ctttatgaac tccctctgtg 1500

tcactgtatg tgaaaggaaa tgcaccaaca accgtaaact gaacgtgttc ttttgtgctc 1560

ttttataact tgcattacat gttgtaagca tggtccgttc tatacctttt tctggtcata 1620

atgaacactc attttgttag cgagggtggt aaagtgaaca aaaaggggaa gtatcaaact 1680

actgccattt cagtgagaaa atcctaggtg ctactttata ataagacatt tgttaggcca 1740

ttcttgcatt gatataaaga aatacctgag actgggtgat ttatatgaaa agaggtttaa 1800

ttggctcaca gttctgcagg ctgtatggga agcatggcgg catctgcttc tggggacacc 1860

tcaggagctt tactcatggc agaaggcaaa gcaaaggcag gcacttcaca cagtaaaagc 1920

aggagcgaga gagaggtgcc acactgaaac agccagatct catgagaagt cactcactat 1980

tgcaaggaca gcatcaaaga gatggtgcta aaccattcat gatgaactca cccccatgat 2040

ccaatcacct cccaccaggc tccacctcga atactgggga ttaccattca gcatgagatt 2100

tgggcaggaa cacagaccca aaccatacca cacacattat cattgttaaa ctttgtaaag 2160

tatttaaggt acatggaaca cacgggaagt ctggtagctc agcccatttc tttattgcat 2220

ctgttattca ccatgtaatt caggtaccac gtattccagg gagcctttct tggccctcag 2280

tttgcagtat acacactttc caagtactct tgtagcatcc tgtttgtatc atagcactgg 2340

tcacattgcc ttacctaaat ctgtttgaca gtctgctcaa cacgactgca agctccatga 2400

gggcagggac atcatctctt ccatctttgg gtccttagtg caatacctgg cagctagcca 2460

gtgctcagct aaatatttgt tgactgaata aatgaatgca caaccaaaaa aaaaaaaaaa 2520

aaaaaaaaaa aaaaaaaaaa aataaaaaaa aaaaaaaaaa aaaaaaaaaa aaa 2573

<210> 51
<211> 803
<212> DNA
<213> Homo Sapiens

<400> 51
ccggcacgag aggagttgtg agtttccaag ccccagctca ctctgaccac ttctctgcct 60

gcccagcatc atgaagggcc ttgcagctgc cctccttgtc ctcgtctgca ccatggccct 120

ctgctcctgt gcacaagttg gtaccaacaa agagctctgc tgcctcgtct atacctcctg 180

gcagattcca caaaagttca tagttgacta ttctgaaacc agcccccagt gccccaagcc 240

aggtgtcatc ctcctaacca agagaggccg gcagatctgt gctgacccca ataagaagtg 300

ggtccagaaa tacatcagcg acctgaagct gaatgcctga ggggcctgga agctgcgagg 360

gcccagtgaa cttggtgggc ccaggaggga acaggagcct gagccagggc aatggccctg 420

ccaccctgga ggccacctct tctaagagtc ccatctgcta tgcccagcca cattaactaa 480

ctttaatctt agtttatgca tcatatttca ttttgaaatt gatttctatt gttgagctgc 540

attatgaaat tagtattttc tctgacatct catgacattg tctttatcat cctttccct 600

ttcccttcaa ctcttcgtac attcaatgca tggatcaatc agtgtgatta gctttctcag 660

cagacattgt gccatatgta tcaaatgaca aatctttatt gaatggtttt gctcagcacc 720

acctttaat atattggcag tacttattat ataaaaggta aaccagcatt ctcactgtga 780

aaaaaaaaaa aaaaaaaaaa aaa 803

<210> 52
<211> 5855
<212> DNA
<213> Homo Sapiens

<400> 52
cacgcagtcc cggcccgagc cgacgccttg caggagggtt caaatccgcg cgggggagct 60

gcgacgcgca agggctgcgg agccgcgggc cggcgagcgc gtcgccacca tgaagcagct 120

gcctgtaccc tgttgaaact tcatggccac agccccaggc cctgctggca ttgccatggg 180

cagcgtgggc agcctgttgg aacggcagga cttctccct gaagagctgc gggcggcact 240

tgccgggtct cggggctccc gccagcctga tgggctcctc cggaagggct tgggccagcg 300

tgagttcctc agctacctgc acctccccaa gaaggacagc aagagcacca agaacaccaa 360

gcgggccct cggaacgagc ctgccgacta tgccaccctc tactaccggg aacattctcg 420

cgcgggtgac ttcagcaaga cctcgctgcc agaacggggt cgctttgaca agtgccgcat 480

tcgcccctca gtgttcaagc ctacggcggg caacgggaaa ggcttcctat ccatgcaaag 540

cctggcgtcc cacaaaggcc agaagctgtg gcgcagcaat ggcagcctgc acacgctggc 600

ctgccacccg cccctgagcc ccgggccccg ggccagccag gcccgggcac agctgctgca 660

cgccctcagc ctagatgagg gcggccctga gcccgagccc agcctgtccg actcctccag 720

tgggggtagt tttggtcgca gtcctggtac tggccctagc cccttcagct cctcccttgg 780

ccaccttaac cacctcgggg gctccctgga ccgggcctct caaggaccca aggaggctgg 840

gccaccagct gtgctgagct gcctgcccga gccaccaccc ccctacgagt tctcctgctc 900

ctctgccgag gaaatgggag ccgtgctgcc cgagacctgt gaggagctca agaggggcct 960

tggcgatgag gacggctcca acccttcac gcaggtgctg gaggagcgcc agcggctgtg 1020

gctggctgag ctgaagcgcc tgtatgtgga gcggctgcac gaggtgaccc agaaggctga 1080

gcgcagcgag cgcaacctcc agctgcagct gtttatggct cagcaggagc agcggcgcct 1140

gcgcaaggag ctgcgggctc agcagggcct ggctccggag cctcgggccc ccggcaccct 1200

cccagaggct gaccccagtg cacgaccaga ggaggaagcc cgatgggagg tgtgccagaa 1260

gacagcagag attagcctct tgaagcagca gctgcgtgaa gcccaggcgg aactggccca 1320

gaagctggcg gagatcttca gtctgaagac acaacttcgg ggcagccggg cacaagccca 1380

ggctcaggac gcagagctgg tccggctgcg cgaggctgtg cgcagcctgc aggagcaggc 1440

ccctcgggag gaagccccag gcagctgtga gactgatgac tgcaagagca ggggcctgct 1500

aggggaggca ggaggcagcg aggccagaga cagtgctgag cagctgcggg ctgagctgct 1560

gcaggagcga cttcggggcc aggagcaggc gctgcgcttt gagcaggagc ggcggacttg 1620

gcaggaggag aaggagcgcg tgctgcgcta ccagcgggag atccagggag ggtacatgga 1680

catgtaccgc cgcaaccagg cactggagca ggaactgcgg gcactgcggg agccccccac 1740

accctggagt ccccggctcg agtcctccaa gatctgaggc cagcagagcg agctgacagc 1800

agcaacactg tcagaaggtg ccctgagacg gccggctcag ccttcccttg cactggttgg 1860

ggtggaacct gcagaggcca gcccggggct ggggaggcgc aaggagagga gggatccagt 1920

ggggccgtgg gctgggtagg gtgccttggc aggagccagg acaaggccct cctggcagag 1980

gagcacctag gcagggccca gccctgcttc ctggagtgga tgtggcccag agaaggaggc 2040

tgggggatca ccagccccaa ggtcccgaag ggcaggtcag agggagagag gctggagacc 2100

tgggctggag ccttcctcca gggaaggagg ctggggtggg aacactggcc tcccccagaa 2160

taaaaccatg ttttctacca gaggctcaga atacgctgag cctgtgacca gaggatgatg 2220

gatggtcggg attgaggctg ttgacctggg cagtagctcc tcccatggcc agtggtcagt 2280

gggagggtgt gccctgcgcc tgtctgcatg gccactgggc atgtgtgttg ggagcagagg 2340

agtcctactc cttgcctcag ccccacacgg ttccttagct gccgtgtggg ctgaaattcc 2400

tttctttagc accagtggag ttttcagaag gaacaacacc agggaatgcc aaaaaaacaa 2460

agggcaagtc aaccaaggca ttttgagaac atgaaatgtc ttcttggtgg gaaggctggg 2520

ccccaggagt atccacaggc tcaggccatg cccctccccg cacccacctc ctgtctgcct 2580

ccctccctcc ataggagtgg ggggccccta gaagtgctct gcctcattcc ctgttcgtag 2640

gaaggtgcag gagaggaggt gggcaagctt atgggtttct tggtaaagag cctttaccac 2700

agcaaggaat gggaacattt ccccatcagc aacggggctc tagggcatta ttaagtaggg 2760

gtgaaatatg attgatttgc attctggaaa gctctcccag gaggaagcat tccccacccc 2820

accttgcctc tgtcctgcgc tgggctccag gacggtcagt cctcccagat ccctgctctc 2880

agctaaggct ggtccctaaa acccacacct gcctttggtc tctcagaggc ctggcttgcc 2940

ctctgggcct ttgctccccc gctgggtgcc cggcctggaa cacaggttct gtctggctgt 3000

gtgagtggcg tctctccctt cctttgaggg aagctcagct ttcttacgcg ctcatcgttg 3060

gacgggctgc catgcattgc gtgcttactg ggcgccaggt gctgtgctgg gcccttgtga 3120

gtgtgggggc acaggctgaa ttagatgagc tcctgccgca gggggccctg cacactgatg 3180

tcagctcacc agtctcttcc ctgacagggc agctcaggga acttctgaga agttatctgc 3240

agagagcccc ccctcactcc ctttggctac cctggccgtg cctggcagcc cctcaggccc 3300

cacaggtggg tcctggttgt gaagtggagt tagtgcacat ggatgtcagg ggacgggtgt 3360

gaggcatgag ctgggggggc ccccttttte agggcagaga atgaggcgtg tgcactgtct 3420

gctgtgggga tggtgtgttt atggccatgt aagtggcagg tgtgtgtaaa gatgtgcttg 3480

tgggtggtgg gtgtgtctga gggggtgtgt ttgggagcct ggcctccaaa agccagtgtc 3540

tgggttggag acagtgcagg ccccttcttg gccccaccct ctaccccatt ccttcctgtc 3600

agccccagtc tgtcttggca ctggccaggc ttgggtggag cctgtgtccc agacagatcc 3660

cctaggccta tctgagaccc atgcaggccc acacctgagc tcctgtctgc ctgaggaggg 3720

aaaggggggag tggccaggtc aggcaggcca gtcgtctaca cacccccttg gtacccattc 3780

ccatttcaca aactcctcca aaccccaaac cccaaacgga aatctctggg acaatggctc 3840

ctccagaggg cccctcagag ccaggtgctg gccagatgcc ttgatcacag cctccatggc 3900

caggggcctg tcaagccagg ggcccacata cctggtgcca gggcccacct cggtatacca 3960

gctcggcaga acagttcccc acccagcgat tagggagctg ggctggccag tacagccagt 4020

ggggagtggg tgggaggcac caaaacttgg cagctgtgcc agaaaccaaa gccaggccca 4080

gatgtagggg gagggcacga gagctgaacc tgagcctggg gcgtctcgga gctggatccc 4140

cgtcagctca tccaggctgg acaaaccagt tcctgggtcc cccagccggc tgggtggctc 4200

tggggccaac cgaccacagg tgtcactgat cacaggtgtc accgtagtgg ctaccacctc 4260

agggcagccc ctccctccag ctggcggggc ctcagcttcc cagctgccct ggcttagcct 4320

tccctgcagc ccctccccaa caggcttggt ccgggcagcc tgctctgcgg aaggtggcag 4380

cctcctggtg ctgggctccc tgcctctggg gcctccagtc cttcactcag tgtcatccct 4440

tccagcagct ccctgggcct cttggtccat cccttcaggc ttctgacaca ctagtggcct 4500

cgagtcactg atagtggtgc ctcctcctta ccttggcaag cccttgtagc cctctgagcc 4560

ctcatcctgc aggcagagcc cctcagtgct ctgagaaacc tgccctgaac tcaccgcctc 4620

ccctgcctgg gcactccctc ctccctgcgc ccacagcccc catgccaatc tccgtcagct 4680

ctttctgtgc tgctgtaatt gttggtttgc ctgtctcctt tgagctgtgg gcattgccgg 4740

ggtagggtgg tgtcttcacc acagctccag aaggcactgg tgggatgtgg agtcaggagc 4800

aaccagagac cccccaatac tagaatgggt ttgagcttgg ggtggggtgg atggggaaga 4860

cttactctga atgttcgttc accatccgtg ggcaccaggt ctgtgccaag cacagggact 4920

ccaggggcag ctgccattcg ttccagtgat gtatttgggg ccttctcagg tgaggaagcc 4980

aaggtggcca aggccctcgg tcccctgccc tgatagggct tacctgtggg tgggggaggg 5040

agaccacaca tgtcagaggt agtgaaggcc gctccaaata gtgactgggg ccaggcaggc 5100

ggaaaagaga acagaagaga atccagaagt gttcagacag aactagggac aacagagggg 5160

cctccatggt ggcatgggtc agcagagcat ggagcaggca gaggaagcct atctgtgggg 5220

ctgggtacat ttctctgaga ctcacggaat gtaagtgttg aggtttctgc aaagagggcg 5280

caggccctgc agtgacggtt cagatattga taacccatcc cccttggcag gtgggggctt 5340

aggcatccat attggattca gaaaagaaaa agaatgatgc ttcttgcaac aaaaacaaaa 5400

gtgattatgt cacgagagga gaggctggag ggccatggga gttcccagct ggctcactca 5460

tccgtccagg aggtgggtct catggtttca ttttgctttg ggcttactta gagttttggc 5520

aaatgcagag ttatgtaacg aacccaacaa tgaagataca gttctgtcac ccccccagca 5580

ttctcctgtg gccctctagt ccacccccag ccccacccca gcccctggca accactgatc 5640

tgctttcaca ataaccgtat ggggtttttt ctctttcctt ttttgtgagc ggtatttaat 5700

tcctttgctc taatatgtaa tattagaaca tattacaaag tagcaagaaa aatacgacat 5760

tggtgtaaaa ataatctaca cctcatggaa agaatgaaga gccagaaata gattctattc 5820

tgtagaaaat cttacacaat aaaggacttc aaaat 5855

<210> 53
<211> 2022
<212> DNA
<213> Homo Sapiens

<400> 53
atgcctggcc agaagttctt cctagaggtt ctatgctgtc ctagcaagaa ttggcgatcg 60

agcgccgcgg agcgcgtccc tccctcgcca atccggctcc ggcgccggcg cccgcccgcg 120

ttttcccggc gcctgccgct ccgccgctcc gacccggcac gcagtcccgg cccgagccga 180

cgccttgcag gagggttcaa atccgcgcgg gggagctgcg acgcgcaagg gctgcggagc 240

cgcgggccgg cgagcgcgtc gccaccatgg gcagctgtgt cttccatctc caccaaggat 300

tggtccgaaa gcaattcctc tccgtgttca gagattccag tgctgcctgc taatcttggg 360

gactggagag ggatttggtg gggaacatgg caggaggccc caggccctgc tggcattgcc 420

atgggcagcg tgggcagcct gttggaacgg caggacttct cccctgaaga gctgcgggcg 480

gcacttgccg ggtctcgggg ctcccgccag cctgatgggc tcctccggaa gggcttgggc 540

cagcgtgagt tcctcagcta cctgcacctc cccaagaagg acagcaagag caccaagaac 600

accaagcggg cccctcggaa cgagcctgcc gactatgcca ccctctacta ccgggaacat 660

tctcgcgcgg gtgacttcag caagacctcg ctgccagaac ggggtcgctt tgacaagtgc 720

cgcattcgcc cctcagtgtt caagcctacg gcgggcaacg ggaaaggctt cctatccatg 780

caaagcctgg cgtcccacaa aggccagaag ctgtggcgca gcaatggcag cctgcacacg 840

ctggcctgcc acccgccccc gagccccggg ccccgggcca gccaggcccg ggcacagctg 900

ctgcacgccc tcagcctaga tgagggcggc cctgagcccg agcccagcct gtccgactcc 960

tccagtgggg gtagttttgg tcgcagtcct ggtactggcc ctagcccctt cagctcctcc 1020

cttggccacc ttaaccacct cgggggctcc ctggaccggg cctctcaagg acccaaggag 1080

gctgggccac cagctgtgct gagctgcctg cccgagccac cacccccta cgagttctcc 1140

tgctcctctg ccgaggaaat gggagccgtg ctgcccgaga cctgtgagga gctcaagagg 1200

ggccttggcg atgaggacgg ctccaacccc ttcacgcagg tgctggagga gcgccagcgg 1260

ctgtggctgg ctgagctgaa gcgcctgtat gtggagcggc tgcacgaggt gacccagaag 1320

gctgagcgca gcgagcgcaa cctccagctg cagctgttta tggctcagca ggagcagcgg 1380

cgcctgcgca aggagctgcg ggctcagcag ggcctggctc cggagcctcg ggcccccggc 1440

accctcccag aggctgaccc cagtgcacga ccagaggagg aagcccgatg ggaggtgtgc 1500

cagaagacag cagagattag cctcttgaag cagcagctgc gtgaagccca ggcggaactg 1560

gcccagaagc tggcggagat cttcagtctg aagacacaac ttcggggcag ccgggcacaa 1620

gcccaggctc aggacgcaga gctggtccgg ctgcgcgagg ctgtgcgcag cctgcaggag 1680

caggcccctc gggaggaagc cccaggcagc tgtgagactg atgactgcaa gagcaggggc 1740

ctgctagggg aggcaggagg cagcgaggcc agagacagtg ctgagcagct gcgggctgag 1800

ctgctgcagg agcgacttcg gggccaggag caggcgctgc gctttgagca ggagcggcgg 1860

acttggcagg aggagaagga gcgcgtgctg cgctaccagc gggagatcca gggagggtac 1920

atggacatgt accgccgcaa ccaggcactg gagcaggaac tgcgggcact gcgggagccc 1980

cccacaccct ggagtccccg gctcgagtcc tccaagatct ga 2022

<210> 54
<211> 3805
<212> DNA
<213> Homo Sapiens

<400> 54
ggctcggctc ctagagctgc cacggccatg gccagagccc gcccgccgcc gccgccgtcg 60
ccgccgccgg ggcttctgcc gctgctccct ccgctgctgc tgctgccgct gctgctgctg 120
cccgccggct gccgggcgct ggaagagacc ctcatggaca caaaatgggt aacatctgag 180
ttggcgtgga catctcatcc agaaagtggg tgggaagagg tgagtggcta cgatgaggcc 240
atgaatccca tccgcacata ccaggtgtgt aatgtgcgcg agtcaagcca gaacaactgg 300
cttcgcacgg ggttcatctg gcggcgggat gtgcagcggg tctacgtgga gctcaagttc 360
actgtgcgtg actgcaacag catccccaac atccccggct cctgcaagga gaccttcaac 420
ctcttctact acgaggctga cagcgatgtg gcctcagcct cctccccctt ctggatggag 480
aaccccctacg tgaaagtgga caccattgca cccgatgaga gcttctcgcg gctggatgcc 540
ggccgtgtca acaccaaggt gcgcagcttt gggccacttt ccaaggctgg cttctacctg 600
gccttccagg accagggcgc ctgcatgtcg ctcatctccg tgcgcgcctt ctacaagaag 660
tgtgcatcca ccaccgcagg cttcgcactc ttccccgaga ccctcactgg ggcggagccc 720
acctcgctgg tcattgctcc tggcacctgc atccctaacg ccgtggaggt gtcggtgcca 780
ctcaagctct actgcaacgg cgatggggag tggatggtgc ctgtgggtgc ctgcacctgt 840
gccaccggcc atgagccagc tgccaaggag tcccagtgcc gcccctgtcc ccctgggagc 900
tacaaggcga agcagggaga ggggccctgc ctcccatgtc cccccaacag ccgtaccacc 960
tccccagccg ccagcatctg cacctgccac aataacttct accgtgcaga ctcggactct 1020
gcggacagtg cctgtaccac cgtgccatct ccaccccgag gtgtgatctc caatgtgaat 1080
gaaacctcac tgatcctcga gtggagtgag ccccgggacc tgggtggccg ggatgacctc 1140
ctgtacaatg tcatctgcaa gaagtgccat ggggctggag gggcctcagc ctgctcacgc 1200
tgtgatgaca acgtggagtt tgtgcctcgg cagctgggcc tgacggagcg ccgggtccac 1260
atcagccatc tgctggccca cacgcgctac acctttgagg tgcaggcggt caacggtgtc 1320
tcgggcaaga gccctctgcc gcctcgttat gcggccgtga atatcaccac aaaccaggct 1380
gccccgtctg aagtgcccac actacgcctg cacagcagct caggcagcag cctcacccta 1440
tcctgggcac ccccagagcg gcccaacgga gtcatcctgg actacgagat gaagtacttt 1500
gagaagagcg agggcatcgc ctccacagtg accagccaga tgaactccgt gcagctggac 1560
gggcttcggc ctgacgcccg ctatgtggtc caggtccgtg cccgcacagt agctggctat 1620
gggcagtaca gccgccctgc cgagtttgag accacaagtg agagaggctc tggggcccag 1680
cagctccagg agcagcttcc cctcatcgtg ggctccgcta cagctgggct tgtcttcgtg 1740
gtggctgtcg tggtcatcgc tatcgtctgc ctcaggaagc agcgacacgg ctctgattcg 1800

gagtacacgg agaagctgca gcagtacatt gctcctggaa tgaaggttta tattgaccct 1860

tttacctacg aggaccctaa tgaggctgtt cgggagtttg ccaaggagat cgacgtgtcc 1920

tgcgtcaaga tcgaggaggt gatcggagct ggggaatttg gggaagtgtg ccgtggtcga 1980

ctgaaacagc ctggccgccg agaggtgttt gtggccatca agacgctgaa ggtgggctac 2040

accgagaggc agcggcggga cttcctaagc gaggcctcca tcatgggtca gtttgatcac 2100

cccaatataa tccggctcga gggcgtggtc accaaaagtc ggccagttat gatcctcact 2160

gagttcatgg aaaactgcgc cctggactcc ttcctccggc tcaacgatgg gcagttcacg 2220

gtcatccagc tggtgggcat gttgcggggc attgctgccg gcatgaagta cctgtccgag 2280

atgaactatg tgcaccgcga cctggctgct cgcaacatcc ttgtcaacag caacctggtc 2340

tgcaaagtct cagactttgg cctctcccgc ttcctggagg atgacccctc cgatcctacc 2400

tacaccagtt ccctgggcgg gaagatcccc atccgctgga ctgccccaga ggccatagcc 2460

tatcggaagt tcacttctgc tagtgatgtc tggagctacg gaattgtcat gtgggaggtc 2520

atgagctatg gagagcgacc ctactgggac atgagcaacc aggatgtcat caatgccgtg 2580

gagcaggatt accggctgcc accacccatg gactgtccca cagcactgca ccagctcatg 2640

ctggactgct gggtgcggga ccggaacctc aggcccaaat tctcccagat tgtcaatacc 2700

ctggacaagc tcatccgcaa tgctgccagc ctcaaggtca ttgccagcgc tcagtctggc 2760

atgtcacagc ccctcctgga ccgcacggtc ccagattaca caaccttcac gacagttggt 2820

gattggctgg atgccatcaa gatggggcgg tacaaggaga gcttcgtcag tgcggggttt 2880

gcatcttttg acctggtggc ccagatgacg gcagaagacc tgctccgtat tggggtcacc 2940

ctggccggcc accagaagaa gatcctgagc agtatccagg acatgcggct gcagatgaac 3000

cagacgctgc ctgtgcaggt ctgacaccgg ctcccacggg gaccctgagg accgtgcagg 3060

gatgccaagc agccggctgg actttcggac tcttggactt ttggatgcct ggccttaggc 3120

tgtggcccag aagctggaag tttgggaaag gcccaagctg ggacttctcc aggcctgtgt 3180

tccctcccca ggaagtgcgc cccaaacctc ttcatattga agatggatta ggagaggggg 3240

tgatgacccc tccccaagcc cctcagggcc cagaccttcc tgctctccag caggggatcc 3300

ccacaacctc acacttgtct gttcttcagt gctggaggtc ctggcagggt caggctgggg 3360

taagccgggg ttccacaggg cccagccctg gcaggggtct ggcccccccag gtaggcggag 3420

agcagtccct ccctcaggaa ctggaggagg ggactccagg aatggggaaa tgtgacacca 3480

ccatcctgaa gccagcttgc acctccagtt tgcacaggga tttgtcctgg gggctgaggg 3540

ccctgtcccc acccccgccc ttggtgctgt cataaaaggg caggcagggg caggctgagg 3600

agttgccctt tgcccccccag agactgactc tcagagccag agatgggatg tgtgagtgtg 3660

tgtgtgtgtg tgtgcgcgcg cgcgcgcgtg tgtgtgtgca cgcactggcc tgcacagaga 3720

gcatgggtga gcgtgtaaaa gcttggccct gtgccctaca atggggccag ctgggccgac 3780

agcagaataa aggcaataag atgaa 3805

<210> 55
<211> 1242
<212> DNA
<213> Homo Sapiens

<400> 55
atgggcggga ctacgctggc ctggagcatg gcacgtgatt ccgccggcct ggttgccggg 60

aatctggacc tgagcgagaa gcacgatccc cggccgcccc cgctcttgca tccccctggt 120

cctactgctg tgcttgctgg cgacggttcg ttccggaagt gtgcagagaa gtctacattc 180

ccatgtcaag ctacagctag agagttgact cctctatttg agccatgcca gccgccacac 240

ctggtgggga gagttaaagg ccgagaagtg aacacagctc caaccccact gccttgtaga 300

ccttccggca gacctgtggc aggtggtgga ggtgatgggc cagggggccc ggagccgggc 360

tgggttgatc ctcggacctg gctaagcttc caaggccctc ctggagggcc aggaatcggg 420

ccgggggttg ggccaggctc tgaggtgtgg gggattcccc catgcccccc gccgtatgag 480

ttctgtgggg ggatggcgta ctgtgggccc caggttggag tggggctagt gccccaaggc 540

ggcttggaga cctctcagcc tgagggtgaa gcaggagtcg gggtggagag caactccgat 600

ggggcctccc cggagccctg caccgtcacc cctggtgccg tgaagctgga gaaggagaag 660

ctggagcaaa acccggagga ggcaaggaag gtattcagcc aaacgaccat ctgccgcttt 720

gaggctctgc agcttagctt caagaacatg tgtaagctgc ggcccttgct gcagaagtgg 780

gtggaggaag ctgacaacaa tgaaaatctt caggagatat gcaaagcaga aaccctcgtg 840

caggcccgaa agagaaagcg aaccagtatc gagaaccgag tgagaggcaa cctggagaat 900

ttgttcctgc agtgcccgaa acccacactg cagcagatca gccacatcgc ccagcagctt 960

gggctcgaga aggatgtggt ccgagtgtgg ttctgtaacc ggcgccagaa gggcaagcga 1020

tcaagcagcg actatgcaca acgagaggat tttgaggctg ctgggtctcc tttctcaggg 1080

ggaccagtgt cctttcctct ggccccaggg ccccattttg gtaccccagg ctatgggagc 1140

cctcacttca ctgcactgta ctcctcggtc cctttccctg aggggaagc ctttcccccc 1200

gtctctgtca ccactctggg ctctcccatg cattcaaact ga 1242

<210> 56
<211> 1380
<212> DNA
<213> Homo Sapiens

<400> 56
ctcatttcac caggcccccg gcttggggcg ccttccttcc ccatggcggg acacctggct 60

tcggatttcg ccttctcgcc ccctccaggt ggtggaggtg atgggccagg ggggccggag 120

ccgggctggg ttgatcctcg gacctggcta agcttccaag gccctcctgg agggccagga 180

atcgggccgg gggttgggcc aggctctgag gtgtgggggga ttcccccatg ccccccgccg 240

tatgagttct gtgggggggat ggcgtactgt gggccccagg ttggagtggg gctagtgccc 300

caaggcggct tggagacctc tcagcctgag ggcgaagcag gagtcggggt ggagagcaac 360

tccgatgggg cctccccgga gccctgcacc gtcacccctg gtgccgtgaa gctggagaag 420

gagaagctgg agcaaaaccc ggaggagtcc caggacatca aagctctgca gaaagaactc 480

gagcaatttg ccaagctcct gaagcagaag aggatcaccc tgggatatac acaggccgat 540

gtggggctca ccctgggggt tctatttggg aaggtattca gccaaacgac catctgccgc 600

tttgaggctc tgcagcttag cttcaagaac atgtgtaagc tgcggccctt gctgcagaag 660

tgggtggagg aagctgacaa caatgaaaat cttcaggaga tatgcaaagc agaaaccctc 720

gtgcaggccc gaaagagaaa gcgaaccagt atcgagaacc gagtgagagg caacctggag 780

aatttgttcc tgcagtgccc gaaacccaca ctgcagcaga tcagccacat cgcccagcag 840

cttgggctcg agaaggatgt ggtccgagtg tggttctgta accggcgcca gaagggcaag 900

cgatcaagca gcgactatgc acaacgagag gattttgagg ctgctgggtc tcctttctca 960

gggggaccag tgtcctttcc tctggcccca gggccccatt ttggtacccc aggctatggg 1020

agccctcact tcactgcact gtactcctcg gtcccttcc ctgaggggga agcctttccc 1080

cctgtctccg tcaccactct gggctctccc atgcattcaa actgaggtgc ctgcccttct 1140

aggaatgggg gacaggggga ggggaggagc tagggaaaga aaacctggag tttgtgccag 1200

ggtttttggg attaagttct tcattcacta aggaaggaat tgggaacaca aagggtgggg 1260

gcaggggagt ttggggcaac tggttggagg gaaggtgaag ttcaatgatg ctcttgattt 1320

taatcccaca tcatgtatca ctttttttctt aaataaagaa gcctgggaca cagtagatag 1380

<210> 57
<211> 1855
<212> DNA
<213> Homo Sapiens

<400> 57
tgcattgcac caggatgtct gtgaaatgga cttcagtaat tttgctaata caactgagct 60

tttgctttag ctctgggaat tgtggaaagg tgctggtgtg ggcagcagaa tacagccatt 120

ggatgaatat aaagacaatc ctggatgagc ttattcagag aggtcatgag gtgactgtac 180

tggcatcttc agcttccatt ctttttgatc ccaacaactc atccgctctt aaaattgaaa 240

tttatcccac atctttaact aaaactgagt tggagaattt catcatgcaa cagattaaga 300

gatggtcaga ccttccaaaa gatacatttt ggttatattt ttcacaagta caggaaatca 360

tgtcaatatt tggtgacata actagaaagt tctgtaaaga tgtagtttca aataagaaat 420

ttatgaaaaa agtacaagag tcaagatttg acgtcatttt tgcagatgct atttttccct 480

gtagtgagct gctggctgag ctatttaaca tacccttttgt gtacagtctc agcttctctc 540

ctggctacac ttttgaaaag catagtggag gatttatttt ccctccttcc tacgtacctg 600

ttgttatgtc agaattaact gatcaaatga ctttcatgga gagggtaaaa aatatgatct 660

atgtgcttta ctttgacttt tggttcgaaa tatttgacat gaagaagtgg gatcagtttt 720

atagtgaagt tctaggaaga cccactacgt tatctgagac aatggggaaa gctgacgtat 780

ggcttattcg aaactcctgg aattttcagt ttcctcatcc actcttacca aatgttgatt 840

ttgttggagg actccactgc aaacctgcca aacccctgcc taaggaaatg gaagactttg 900

tacagagctc tggagaaaat ggtgttgtgg tgttttctct ggggtcaatg gtcagtaaca 960

tgacagaaga aagggccaac gtaattgcat cagccctggc ccagatccca caaaaggttc 1020

tgtggagatt tgatgggaat aaaccagata ccttaggtct caatactcgg ctgtataagt 1080

ggatacccca gaatgacctt ctaggtcatc caaagaccag agcttttata actcatggtg 1140

gagccaatgg catctacgag gcaatctacc atgggatccc tatggtgggg attccattgt 1200

ttgccgatca acctgataac attgctcaca tgaaggccag gggagcagct gttagagtgg 1260

acttcaacac aatgtcgagt acagacttgc tgaatgcatt gaagagagta attaatgatc 1320

cttcatataa agagaatgtt atgaaattat caagaattca acatgatcaa ccagtgaagc 1380

ccctggatcg agcagtcttc tggattgaat ttgtcatgcg ccacaaagga gctaaacacc 1440

ttcgggttgc agcccacgac ctcacctggt tccagtacca ctctttggat gtgattgggt 1500

tcctgctggt ctgtgtggca actgtgatat ttatcgtcac aaaatgttgt ctgttttgtt 1560

tctggaagtt tgctagaaaa gcaaagaagg gaaaaaatga ttagttatat ctgagatttg 1620

aagctggaaa acctgatagg tgagactact tcagtttatt ccagcaagaa agattgtgat 1680

gcaagatttc tttcttcctg agacaaaaaa aaaaaaaaga aaaaaaaatc ttttcaaaat 1740

ttactttgtc aaataaaaat ttgtttttca gagatttacc acccagttca tggttagaaa 1800

tattttgtgg caatgaagaa aacactacgg aaaataaaaa ataagataaa gcctt 1855

<210> 58
<211> 8619
<212> DNA
<213> Homo Sapiens

<400> 58
atgcttcagt gtacaccagc caatatggta gaagttcaca aagacaaaga gtcaagcaaa 60
ggtcacacta gacacaaagt ggaagaagct cttattaatg aagaagcaat tttgaacctt 120
atggaaaata gtcagacttt tcagcctttg acccaaagac tgagtgagtc acctgttttc 180
atggacagta gtcctgatga ggctctggta catcttcttg ctggtttgga aagtgatgga 240
tatcgggggg aaagaaatag gatgccatca ccatgtcgct cctttggaaa taataaatat 300
ccacaaaata gtgatgatga agaaaatgaa ccacagattg aaaaagagga aatggagctt 360
agtttggtga tgtcccagag atgggacagc aatattgaag aacattgtgc caaaaagaga 420
tcactgtgca gaaataccca cagaagttca actgaagatg atgactcatc ttcaggagaa 480
gaaatggaat ggagtgataa cagtttgctt ctagccagtc tttctatacc tcagttagat 540
ggaactgcag atgaaaatag tgacaatcca ttgaacaatg aaaattctag aacccactct 600
tctgtaattg caacaagcaa gctttcagtt aaaccctcca tctttcacaa agatgctgct 660
acattagaac cctcatcttc tgctaagatt acctttcagt gtaaacacac aagtgccctt 720
tcttcccatg ttttgaacaa ggaagattta attgaagacc tttcacagac aaacaaaaat 780
acagaaaaag gtctagataa ctcagtcact tcttttacaa acgaaagcac ttattctatg 840
aaataccctg gatctttaag cagtactgtt cattcagaaa attctcataa agagaatagt 900
aagaaagaga tcctcccagt atcttcctgt gaaagtagta tttttgatta tgaagaagat 960
attccatctg ttacaagaca agtaccaagt agaaaatata caaacattag aaaaatcgaa 1020
aaggattccc cttttataca tatgcaccgt caccctaacg agaatacatt gggcaaaaat 1080
tctttcaact tttctgactt aaatcattca aaaaataaag tatcctctga aggaaatgaa 1140
aaaggaaaca gcacagctct gagtagttta ttcccttcat catttactga aaattgtgaa 1200
ttactgtcat gctcagggga gaatagaact atggtgcatt ctcttaatag cactgctgat 1260
gaaagtggac taaataaact taaaattagg tatgaagaat ttcaagaaca taaaacagaa 1320
aagccaagcc tcagccagca agcagcacac tatatgtttt ttcccagtgt tgttctttct 1380
aactgtctta ctagaccaca gaaactatct cctgtcacat ataaattaca acctggcaat 1440
aaaccatccc ggttaaaatt gaataaaagg aaacttgcag gtcatcagga gacttctacc 1500
aaaagtagtg agactggatc cacaaaagat aattttatac aaaataatcc ttgtaatagt 1560
aatcctgaga aggataatgc attggctagt gatttaacta aaaccactcg tggagctttt 1620
gaaaataaaa cacccacaga tggttttata gactgtcact ttggagatgg aacgttagaa 1680
actgagcagt cctttggact atatggaaat aaatacacac ttagagccaa acgcaaggta 1740
aattatgaga ctgaagacag tgagtcaagt tttgtaactc acaactcaaa aattagtcta 1800

cctcatccca tggaaattgg tgaaagttta gatggaactc tcaaatcccg aaaacgaaga 1860

aaaatgtcta aaaagctgcc ccctgtcatc ataaagtata ttattattaa tagatttaga 1920

gggagaaaaa atatgcttgt gaagctagga aaaatagact ctaaagaaaa acaagtaata 1980

ttaacagaag aaaaaatgga actatataaa aagcttgcac ctttgaagga cttttggcca 2040

aaagttcccg actcccctgc aaccaaatat cccatttatc cactaacacc aaagaaaagt 2100

cacagaagaa agtcaaaaca taaatctgct aagaaaaaaa ctggtaaaca acaaaggaca 2160

aataatgaaa atattaaaag aactttgtct ttcaggaaaa aacggtcaca tgctattctt 2220

tctcctccct caccatctta caatgctgaa accgaagatt gtgacctgaa ttatagtgat 2280

gttatgtcta aactaggttt tctttctgag agaagcacaa gtcccataaa ttcttctcca 2340

cctcgctgct ggtctcccac agatccaaga gctgaagaaa tcatggctgc tgcagaaaaa 2400

gaggcaatgc tttttaaggg tcctaatgta tataagaaga ctgttaattc tcgtatagga 2460

aaaactagtc gcgcaagagc acagattaag aaatcaaaag caaagcttgc taatccctct 2520

atagttacta agaaaaggaa caaacgaaat cagacaaata aactagtaga tgatggaaaa 2580

aagaaaccaa gagcaaaaca aaaaacaaat gagaaaggta catcgagaaa gcatacaaca 2640

cttaaggatg aaaaaataaa atctcagtct ggtgctgagg ttaagtttgt actgaaacac 2700

cagaatgtgt ctgaatttgc aagtagttct ggaggctctc aactactttt taaacagaaa 2760

gatatgccac taatgggctc tgctgtagat catccccttt ctgcttccct acccactgga 2820

attaatgcac aacagaagtt atctggctgc ttttcttctt tcttagaaag caagaagtct 2880

gtagatttgc agacattccc cagttcacga gatgatttgc atccatcagt tgtttgtaat 2940

tctataggac ctggagtctc aaaaattaat gttcaaaggc ctcataatca aagtgctatg 3000

tttactctaa aggaatcaac gttaattcaa aaaaatatat ttgacctttc caatcattta 3060

tctcaggtag cacagaatac acagatatct tctggtatgt cctcaaagat agaagataat 3120

gcaaataata tacaaagaaa ctatttgtca tcaatcggaa agttaagtga atatcgcaat 3180

tccctagaat caaagctgga ccaagcatat acccctaatt ttttgcattg caaagacagt 3240

cagcagcaga ttgtgtgcat agcggaacag tcaaagcaca gtgaaacttg ttctccggga 3300

aatacagctt cagaggaaag ccaaatgcct aataattgct ttgtaacttc cttgagaagt 3360

ccaatcaaac aaatagcatg ggagcaaaag caaaggggct ttattttaga tatgtcaaat 3420

tttaaacctg aaagagtaaa accgaggtcg ttatcagaag caatttcaca aaccaaagca 3480

ctttctcagt gtaaaaatcg aaatgtgtca acaccttcag catttggtga aggacagtct 3540

ggactggcag ttctaaaaga attgttacaa aaaagacagc agaaagcaca aaatgcaaat 3600

actacacaag acccattatc caataaacat caaccaaata aaaatatttc tggttccctt 3660

gagcataaca aagcaaataa acggacacga tcggtaacgt ccccaagaaa acctcgaact 3720

cccagaagta caaaacaaaa agaaaaaatc cccaaacttc tcaaagtaga ctctttaaat 3780

ttacaaaact ctagccagtt ggataactct gtatcagatg atagtcccat ctttttttca 3840

gatccaggct ttgaaagttg ttactcactt gaagatagtt tatctcctga acataattat 3900

aattttgata ttaacacaat aggtcagact ggattttgta gcttttattc tggaagtcag 3960

tttgtcccag ctgatcagaa tttgcctcag aagttcctaa gtgatgctgt tcaggatctt 4020

tttccaggac aagctataga aaaaaatgag tttttaagtc atgacaacca gaaatgtgat 4080

gaagacaagc atcataccac agactcagcc tcatggatta gatctggtac tttaagtcct 4140

gaaatttttg agaagtcaac catagatagc aatgagaatc gtcgccacaa ccagtggaaa 4200

aatagctttc atcctctaac aactcggtct aactcaataa tggattcttt ctgtgttcag 4260

caggcagaag actgtctaag tgaaaaatct agattgaata ggagttcagt aagcaaagaa 4320

gtgtttctta gcctcccaca gccaaacaat tcagactgga ttcaaggtca caccagaaaa 4380

gaaatgggac agtctcttga ctcagccaat acctctttta ctgcaatact ctcctccct 4440

gatggtgaac ttgtagacgt ggcctgtgaa gatttagaac tgtatgtttc aagaaacaat 4500

gatatgttga caccaactcc tgatagttca ccaagatcta ctagctctcc ttcacaatct 4560

aaaaatggca gcttcacccc tcgaactgct aacattctga aaccacttat gtcccccca 4620

agtagggaag aaattatggc aactttgttg gatcatgacc tgtctgagac tatttaccag 4680

gaaccatttt gcagtaatcc ttctgatgta ccagaaaagc ccagggagat tggtggacgg 4740

ctcctcatgg tagaaactcg acttgcaaat gatctggctg agtttgaggg agacttttcc 4800

ttggaaggac ttcgtctttg gaaaacagca ttctcagcaa tgactcagaa tccaaggcca 4860

gggtcacccc ttcgcagtgg ccaaggagtt gtcaataaag ggtcaagtaa tagccctaag 4920

atggttgaag ataaaaaaat tgtgattatg ccttgcaaat gtgccccaag tcgacaactg 4980

gttcaagtgt ggcttcaagc caaagaagaa tacgaacgtt ccaagaaact gcctaaaacc 5040

aagccaactg gagttgtaaa atctgctgag aactttagct cttcagttaa cccagatgac 5100

aaacctgtag tgcctccaaa aatggatgta agtccatgta tactccccac tacagcacat 5160

accaaggagg atgttgataa ttctcagatt gctttacaag caccaaccac gggatgtagt 5220

caaactgcaa gtgaaagtca gatgctgcca ccagttgcct ctgcaagtga tcccgaaaaa 5280

gatgaagatg atgatgataa ctattacatt agttatagct cccctgattc tccagtaatt 5340

ccccccttggc aacaaccaat atccccagat tccaaagcat taaatggaga tgatagaccc 5400

tcatcaccag tagaggagct gccttcattg gcttttgaga acttcttaaa gccaataaaa 5460

gatggtatac aaaaaagccc ctgcagtgag cctcaagagc ctctagtgat atctccaatt 5520
aatactaggg caagaactgg gaaatgtgaa tcactttgct ttcatagtac accaatcata 5580
cagagaaaac ttctggaaag gcttcctgaa gcacctggcc ttagcccatt atcaacagaa 5640
ccaaaaacac agaagttgag taataagaaa ggaagtaata ctgacactct tagaagagta 5700
ctgttaacac aagcaaagaa tcaatttgca gcagtaaata ccccacagaa agaaacttct 5760
cagattgatg gaccatcttt aaacaatact tacggtttca aagtcagcat acaaaactta 5820
caggaggcaa aagctttaca tgagatacaa aatcttaccc taatcagtgt ggagttgcat 5880
gctcgaacta gacgagactt agaaccggat cctgaatttg acccaatctg tgctctgttc 5940
tactgcatct catctgacac tccactgcca gatacagaaa aaacagaact cacaggtgta 6000
atagtgattg ataaagacaa gacagttttc agtcaagata tcagatatca gactccatta 6060
cttattagat ctggaattac aggactcgaa gtcacctatg ctgctgatga gaaggcactt 6120
tttcatgaaa ttgcaaatat aataaagagg tatgatcctg atattctgct aggatatgag 6180
attcagatgc attcctgggg ttacctctta caaagggctg ccgctttaag tattgactta 6240
tgtcggatga tctctcgggt gccagatgac aaaattgaga acagatttgc agctgaaaga 6300
gatgagtatg gatcatatac aatgagtgag ataaatattg ttggccgaat tacactaaat 6360
ctttggagaa tcatgagaaa tgaggtggct ctaactaact acacctttga aaatgtgagc 6420
tttcatgttc ttcatcagcg ttttcccctc tttacctttc gagtcttgtc agactggttt 6480
gataacaaga cagatctata caggtactgt tctataactc tgaagaagag gcaacagacc 6540
tctgctttgt accactggca ggtcctgggc ccaatatact tctgggtcat ttttacatct 6600
tataatatta aaattctttt tatggatttg ctgagggttt tattgtttgt tttcttaaga 6660
agatggaaaa tggttgatca ttatgttagc cgtgtccgtg gaaatctcca aatgttagaa 6720
cagctggacc tgattgggaa aaccagtgag atggctagac tttttggcat tcagttttta 6780
catgtactga caaggggttc acagtaccgt gtggaatcaa tgatgttgcg tattgctaaa 6840
ccaatgaact atattcctgt gacacctagt gttcagcaaa gatcccagat gagagcccca 6900
cagtgtgttc ctctaattat ggagcctgaa tcccgcttct atagcaactc tgttctcgtt 6960
ttggatttcc aatcacttta tccttctatt gtgattgcat ataactactg cttttccacc 7020
tgccttggcc atgtggagaa cttgggaaag tatgatgagt tcaaatttgg ctgtacctct 7080
ctgagagtac ctccagattt actttaccaa gttaggcatg atatcacagt gtcccccaat 7140
ggagtagctt ttgtcaagcc ttcagtaaga aaaggtgtac taccaagaat gcttgaagaa 7200
attttgaaga ctagatttat ggtgaagcag tcaatgaagg cttacaagca agacagagcc 7260

ctgtcacgaa tgcttgatgc gcgtcagttg ggacttaagc tgatagcaaa tgtcacattt 7320

ggctatacat ctgctaattt ttctgggaga atgccatgca ttgaggttgg cgatagtatt 7380

gttcacaaag ccagagagac cttggaacga gctattaaac tggtgaatga taccaagaaa 7440

tgggggggcta gggttgtata tggcgatact gacagtatgt ttgtgctact gaaaggagcc 7500

actaaggagc agtcttttaa gattggtcag gaaattgccg aagctgtaac tgctaccaat 7560

cctaaaccag tgaaattgaa gtttgaaaag gtatatttgc cctgtgtttt acaaacaaaa 7620

aagaggtatg tgggttacat gtatgaaaca ctggatcaga aggacccagt atttgatgca 7680

aaaggaatag aaacagtcag aagagattcc tgccctgctg tttctaagat acttgagcgt 7740

tctctaaagc tgctatttga aacgagagat ataagtctaa ttaaacagta tgttcagcga 7800

caatgtatga agcttctgga aggaaaggcc agcatacaag actttatctt tgccaaggaa 7860

tacagaggaa gtttttctta taaaccagga gcttgtgtgc cagcccttga acttacaagt 7920

tttttcattg ttttattatt gtttaattct gacttaattt gtgagaaaga tggcttccat 7980

aacagtattt gggtgtggtt ttttttctttg aattcgaata ggaaaatgct gacttatgac 8040

cggcgctctg agcctcaggt tggggagcga gtgccatacg tcatcattta tgggaccccc 8100

ggagtaccac ttatccagct tgtaaggcgc ccagtggaag tcctgcagga cccaactctg 8160

agactgaatg ctacttacta tattaccaag caaatccttc caccttggc aagaatcttc 8220

tcacttattg gtattgatgt cttcagctgg tatcatgaat taccaaggat ccataaagct 8280

accagctcct cgcgaagtga acctgaaggg cggaaaggca ctatttcaca atattttact 8340

accttacact gtcctgtgtg tgatgaccta actcagcatg gcatctgtag taaatgtcgg 8400

agccaacctc agcatgttgc agtcatcctc aaccaagaaa tccgggagtt ggaacgtcaa 8460

caggagcaac ttgtaaagat atgcaagaac tgtacaggtt gctttgatcg acacatccca 8520

tgtgtttctc tgaactgccc agtacttttc aaactctccc gagtaaatag agaattgtcc 8580

aaggcaccat atctccggca gttattagac cagttttaa 8619

<210> 59
<211> 2335
<212> PRT
<213> Homo Sapiens

<400> 59

Met Ala Met Met Ile Leu Arg Val Asp Tyr Thr Phe Glu Glu Asn Arg
1 5 10 15

Asp Lys Leu Ala Ser Arg Lys Lys Glu Tyr Ser Gln Gly Ser Val Ala
20 25 30

```
Asp Leu Thr Pro Asp Asn Trp Lys Asn Ile Thr Val Pro His Ser Gly
        35                  40                  45

Arg His Ser Glu Val Ser Arg Gly Glu Leu Val Cys Arg Thr Cys Ser
    50                  55                  60

Glu Cys Ser Ala Gly Pro His Ile Trp Met Lys Gly Leu Tyr Gln Thr
65                  70                  75                  80

Gln Asp Glu Glu Ala Gly Gly Glu Asn Ile Phe Ile Leu Leu Phe Ile
                85                  90                  95

Glu Ser Thr Gln Phe Gly Gln Phe Val Ala Met Gly Ser Pro Ile Thr
            100                 105                 110

Glu His Lys Val Phe Thr Met Tyr Leu Gly Leu Ala Thr His Leu Phe
        115                 120                 125

Tyr Ser Leu Ile Thr His Pro Phe Val Leu Leu Glu Asn His Ser Cys
    130                 135                 140

Pro Ser Ser Val His Gly Phe Asp Val Ala Gly Leu Ile Phe Asp Lys
145                 150                 155                 160

Val Gly Met Arg Ser Arg Pro Gly Arg Met Gly Ala Leu Phe Ala Tyr
                165                 170                 175

Phe Ala Gly Phe Ile Arg Arg Lys Ala Leu Val Val Cys Leu Phe Val
            180                 185                 190

Phe Cys Trp Ser Asn Glu Ala Ala Asn Lys Pro Pro Ile Gln Glu Ala
        195                 200                 205

Ala Gln Leu Ser Arg Pro Ala Gln Gly Ala Arg Arg Ala Ser Glu Arg
    210                 215                 220

Lys Phe Leu Ala Phe Ser Cys Pro Leu Ala Gly His Tyr Ala Ala Lys
225                 230                 235                 240

Gln Pro Ser Pro Ser Pro Pro Pro Pro Pro Ala Pro Pro Ala Pro Pro
                245                 250                 255

Ala Ala Arg Ala Ala Gln Leu Ser Ala Gly Gly Gly Val Ala Gln Pro
            260                 265                 270
```

```
Ser Ala Asp Gly Thr Leu Ala Ala Arg Pro Gln Arg Leu Leu Lys Ser
        275                 280                 285

Lys Val Gly Gly Gly Arg Arg Ala Pro Arg Ala Leu His Gly Arg Cys
    290                 295                 300

Leu Ala Ser Pro Pro Gln Pro Arg Arg Ala Gly Gly Arg Gly Val Gly
305                 310                 315                 320

Ala Ala Glu Gly Gly Val Gly Ser Thr Met Gln Phe Val Ser Trp Ala
                325                 330                 335

Thr Leu Leu Thr Leu Leu Val Arg Asp Leu Ala Glu Met Gly Ser Pro
            340                 345                 350

Asp Ala Ala Ala Ala Val Arg Lys Asp Arg Leu His Pro Arg Gln Val
        355                 360                 365

Lys Leu Leu Glu Thr Leu Ser Glu Tyr Glu Ile Val Ser Pro Ile Arg
    370                 375                 380

Val Asn Ala Leu Gly Glu Pro Phe Pro Thr Asn Val His Phe Lys Arg
385                 390                 395                 400

Thr Arg Arg Ser Ile Asn Ser Ala Thr Asp Pro Trp Pro Ala Phe Ala
                405                 410                 415

Ser Ser Ser Ser Ser Ser Thr Ser Ser Gln Ala His Tyr Arg Leu Ser
            420                 425                 430

Ala Phe Gly Gln Gln Phe Leu Phe Asn Leu Thr Ala Asn Ala Gly Phe
        435                 440                 445

Ile Ala Pro Leu Phe Thr Val Thr Leu Leu Gly Thr Pro Gly Val Asn
    450                 455                 460

Gln Thr Lys Phe Tyr Ser Glu Glu Glu Ala Glu Leu Lys His Cys Phe
465                 470                 475                 480

Tyr Lys Gly Tyr Val Asn Thr Asn Ser Glu His Thr Ala Val Ile Ser
                485                 490                 495

Leu Cys Ser Gly Met Gly Leu Leu Asp Val Ser Glu Leu Ser Gly Val
            500                 505                 510

Trp Thr Arg Phe Ser Gly Ala Leu Pro Asn Ala Ala Arg Arg Pro Gly
```

515 520 525

Ser Gln Phe Pro Asn Ser Glu Lys Val Thr Gly Val Ala Val Pro Cys
530 535 540

Ser Lys Leu Gly His Pro Gly Ala Glu Pro Leu Ser Ala Gly Arg Thr
545 550 555 560

Arg Leu Leu Ile Val Asp Leu Thr Arg His Leu Pro Pro Thr Ser Pro
565 570 575

Arg His Leu Arg Ser Arg Cys Gly Thr Val Leu Ala Arg Ala Arg Val
580 585 590

Val Leu Asp Phe Pro Lys Arg Arg Ala Phe Leu Pro Arg Ala Cys Asp
595 600 605

Ala Glu Thr Phe Pro Ala Gly Pro Trp Ile Leu Thr Pro Arg His Trp
610 615 620

Ala Ala Pro Ser Val Arg Cys Arg Ser Trp Val Leu Lys Phe Pro Ser
625 630 635 640

Thr Ser Phe Leu Leu Cys Leu Ser Met Glu Gly Ser Gly Gly Glu Arg
645 650 655

Gly Lys Pro Glu Asp Trp Glu Gly Val Val Leu Ala Cys Trp Asp Ser
660 665 670

Arg Lys Gly Ile Asn Pro Phe Ser Pro Gln Gln Ser Ala Arg Ser Arg
675 680 685

Gly Ser Arg Asn Ala Leu Ser Arg Leu Phe Gly Gly Gly Arg Arg Arg
690 695 700

Gln Leu Gly Glu Val Gly Gly Gly Ala Ala Leu Gly Thr Phe Arg Ser
705 710 715 720

His Asp Gly Asp Tyr Phe Ile Glu Pro Leu Gln Ser Met Asp Glu Gln
725 730 735

Glu Asp Glu Glu Glu Gln Asn Lys Pro His Ile Ile Tyr Arg Arg Ser
740 745 750

Ala Pro Gln Arg Glu Pro Ser Thr Gly Arg His Ala Cys Asp Thr Ser
755 760 765

```
Gly Leu Gln Lys Cys Leu Ile Asn Gly Ser His Glu Asn Ile Tyr Val
    770                 775                 780

Phe Val Glu Cys Phe Leu Glu Thr Ser Gly Leu Leu Met Phe Cys Asp
785                 790                 795                 800

Leu Arg Asn Cys Ser Lys Val Pro Val Arg Tyr Ala Val Ser Tyr Phe
                805                 810                 815

Cys Thr Pro Ser Leu Asn Ser Asp Ala Ala Ser Gln Asn Ser Leu Glu
            820                 825                 830

Tyr Gly Thr Ile His Gln Gln Val Ser Glu Glu Trp Thr Asn Arg Ser
        835                 840                 845

Arg Thr Pro Leu Glu Pro Glu His Lys Asn Arg His Ser Lys Asp Lys
    850                 855                 860

Lys Lys Thr Arg Ala Arg Lys Trp Gly Glu Arg Ile Asn Leu Ala Gly
865                 870                 875                 880

Asp Val Ala Ala Leu Asn Ser Gly Leu Ala Thr Glu Ala Phe Ser Ala
                885                 890                 895

Tyr Gly Asn Lys Thr Asp Asn Thr Arg Glu Lys Arg Thr His Arg Arg
            900                 905                 910

Thr Lys Arg Phe Leu Ser Tyr Pro Arg Phe Val Glu Val Leu Val Val
        915                 920                 925

Ala Asp Asn Arg Met Val Ser Tyr His Gly Glu Asn Leu Gln His Tyr
    930                 935                 940

Ile Leu Thr Leu Met Ser Ile Val Ala Ser Ile Tyr Lys Asp Pro Ser
945                 950                 955                 960

Ile Gly Asn Leu Ile Asn Ile Val Ile Val Asn Leu Ile Val Ile His
                965                 970                 975

Asn Glu Gln Asp Gly Pro Ser Ile Ser Phe Asn Ala Gln Thr Thr Leu
            980                 985                 990

Lys Asn Phe Cys Gln Trp Gln His  Ser Lys Asn Ser Pro  Gly Gly Ile
        995                 1000                1005
```

```
His His  Asp Thr Ala Val Leu  Leu Thr Arg Gln Asp  Ile Cys Arg
  1010                1015                1020

Ala His  Asp Lys Cys Asp Thr  Leu Gly Leu Ala Glu  Leu Gly Thr
    1025                1030                1035

Ile Cys  Asp Pro Tyr Arg Ser  Cys Ser Ile Ser Glu  Asp Ser Gly
    1040                1045                1050

Leu Ser  Thr Ala Phe Thr Ile  Ala His Glu Leu Gly  His Val Phe
    1055                1060                1065

Asn Met  Pro His Asp Asp Asn  Asn Lys Cys Lys Glu  Glu Gly Val
    1070                1075                1080

Lys Ser  Pro Gln His Val Met  Ala Pro Thr Leu Asn  Phe Tyr Thr
    1085                1090                1095

Asn Pro  Trp Met Trp Ser Lys  Cys Ser Arg Lys Tyr  Ile Thr Glu
    1100                1105                1110

Phe Leu  Asp Thr Gly Tyr Gly  Glu Cys Leu Leu Asn  Glu Pro Glu
    1115                1120                1125

Ser Arg  Pro Tyr Pro Leu Pro  Val Gln Leu Pro Gly  Ile Leu Tyr
    1130                1135                1140

Asn Val  Asn Lys Gln Cys Glu  Leu Ile Phe Gly Pro  Gly Ser Gln
    1145                1150                1155

Val Cys  Pro Tyr Met His Cys  Lys Tyr Gly Phe Cys  Val Pro Lys
    1160                1165                1170

Glu Met  Asp Val Pro Val Thr  Asp Gly Ser Trp Gly  Ser Trp Ser
    1175                1180                1185

Pro Phe  Gly Thr Cys Ser Arg  Thr Cys Gly Gly Gly  Ile Lys Thr
    1190                1195                1200

Ala Ile  Arg Glu Cys Asn Arg  Pro Glu Pro Lys Asn  Gly Gly Lys
    1205                1210                1215

Tyr Cys  Val Gly Arg Arg Met  Lys Phe Lys Ser Cys  Asn Thr Glu
    1220                1225                1230
```

```
Pro Cys  Leu Lys Gln Lys Arg  Asp Phe Arg Asp Glu  Gln Cys Ala
    1235                 1240                 1245

His Phe  Asp Gly Lys His Phe  Asn Ile Asn Gly Leu  Leu Pro Asn
    1250                 1255                 1260

Val Arg  Trp Val Pro Lys Tyr  Ser Gly Ile Leu Met  Lys Asp Arg
    1265                 1270                 1275

Cys Lys  Leu Phe Cys Arg Val  Ala Gly Asn Thr Ala  Tyr Tyr Gln
    1280                 1285                 1290

Leu Arg  Asp Arg Val Ile Asp  Gly Thr Pro Cys Gly  Gln Asp Thr
    1295                 1300                 1305

Asn Asp  Ile Cys Val Gln Gly  Leu Cys Arg Gln Ala  Gly Cys Asp
    1310                 1315                 1320

His Val  Leu Asn Ser Lys Ala  Arg Arg Asp Lys Cys  Gly Val Cys
    1325                 1330                 1335

Gly Gly  Asp Asn Ser Ser Cys  Lys Thr Val Ala Gly  Thr Phe Asn
    1340                 1345                 1350

Thr Val  His Tyr Gly Tyr Asn  Thr Val Val Arg Ile  Pro Ala Gly
    1355                 1360                 1365

Ala Thr  Asn Ile Asp Val Arg  Gln His Ser Phe Ser  Gly Glu Thr
    1370                 1375                 1380

Asp Asp  Asp Asn Tyr Leu Ala  Leu Ser Ser Ser Lys  Gly Glu Phe
    1385                 1390                 1395

Leu Leu  Asn Gly Asn Phe Val  Val Thr Met Ala Lys  Arg Glu Ile
    1400                 1405                 1410

Arg Ile  Gly Asn Ala Val Val  Glu Tyr Ser Gly Ser  Glu Thr Ala
    1415                 1420                 1425

Val Glu  Arg Ile Asn Ser Thr  Asp Arg Ile Glu Gln  Glu Leu Leu
    1430                 1435                 1440

Leu Gln  Val Leu Ser Val Gly  Lys Leu Tyr Asn Pro  Asp Val Arg
    1445                 1450                 1455

Tyr Ser  Phe Asn Ile Pro Ile  Glu Asp Lys Pro Gln  Gln Phe Tyr
```

1460 1465 1470

Trp Asn Ser His Gly Pro Trp Gln Ala Cys Ser Lys Pro Cys Gln
1475 1480 1485

Gly Glu Arg Lys Arg Lys Leu Val Cys Thr Arg Glu Ser Asp Gln
1490 1495 1500

Leu Thr Val Ser Asp Gln Arg Cys Asp Arg Leu Pro Gln Pro Gly
1505 1510 1515

His Ile Thr Glu Pro Cys Gly Thr Asp Cys Asp Leu Arg Trp His
1520 1525 1530

Val Ala Ser Arg Ser Glu Cys Ser Ala Gln Cys Gly Leu Gly Tyr
1535 1540 1545

Arg Thr Leu Asp Ile Tyr Cys Ala Lys Tyr Ser Arg Leu Asp Gly
1550 1555 1560

Lys Thr Glu Lys Val Asp Asp Gly Phe Cys Ser Ser His Pro Lys
1565 1570 1575

Pro Ser Asn Arg Glu Lys Cys Ser Gly Glu Cys Asn Thr Gly Gly
1580 1585 1590

Trp Arg Tyr Ser Ala Trp Thr Glu Cys Ser Lys Ser Cys Asp Gly
1595 1600 1605

Gly Thr Gln Arg Arg Arg Ala Ile Cys Val Asn Thr Arg Asn Asp
1610 1615 1620

Val Leu Asp Asp Ser Lys Cys Thr His Gln Glu Lys Val Thr Ile
1625 1630 1635

Gln Arg Cys Ser Glu Phe Pro Cys Pro Gln Trp Lys Ser Gly Asp
1640 1645 1650

Trp Ser Glu Cys Leu Val Thr Cys Gly Lys Gly His Lys His Arg
1655 1660 1665

Gln Val Trp Cys Gln Phe Gly Glu Asp Arg Leu Asn Asp Arg Met
1670 1675 1680

Cys Asp Pro Glu Thr Lys Pro Thr Ser Met Gln Thr Cys Gln Gln
1685 1690 1695

```
Pro Glu  Cys Ala Ser Trp Gln  Ala Gly Pro Trp Gly  Gln Cys Ser
    1700                 1705                 1710

Val Thr  Cys Gly Gln Gly Tyr  Gln Leu Arg Ala Val  Lys Cys Ile
    1715                 1720                 1725

Ile Gly  Thr Tyr Met Ser Val  Val Asp Asp Asn Asp  Cys Asn Ala
    1730                 1735                 1740

Ala Thr  Arg Pro Thr Asp Thr  Gln Asp Cys Glu Leu  Pro Ser Cys
    1745                 1750                 1755

His Pro  Pro Pro Ala Ala Pro  Glu Thr Arg Arg Ser  Thr Tyr Ser
    1760                 1765                 1770

Ala Pro  Arg Thr Gln Trp Arg  Phe Gly Ser Trp Thr  Pro Cys Ser
    1775                 1780                 1785

Ala Thr  Cys Gly Lys Gly Thr  Arg Met Arg Tyr Val  Ser Cys Arg
    1790                 1795                 1800

Asp Glu  Asn Gly Ser Val Ala  Asp Glu Ser Ala Cys  Ala Thr Leu
    1805                 1810                 1815

Pro Arg  Pro Val Ala Lys Glu  Glu Cys Ser Val Thr  Pro Cys Gly
    1820                 1825                 1830

Gln Trp  Lys Ala Leu Asp Trp  Ser Ser Cys Ser Val  Thr Cys Gly
    1835                 1840                 1845

Gln Gly  Arg Ala Thr Arg Gln  Val Met Cys Val Asn  Tyr Ser Asp
    1850                 1855                 1860

His Val  Ile Asp Arg Ser Glu  Cys Asp Gln Asp Tyr  Ile Pro Glu
    1865                 1870                 1875

Thr Asp  Gln Asp Cys Ser Met  Ser Pro Cys Pro Gln  Arg Thr Pro
    1880                 1885                 1890

Asp Ser  Gly Leu Ala Gln His  Pro Phe Gln Asn Glu  Asp Tyr Arg
    1895                 1900                 1905

Pro Arg  Ser Ala Ser Pro Ser  Arg Thr His Val Leu  Gly Gly Asn
    1910                 1915                 1920
```

```
Gln Trp  Arg Thr Gly Pro Trp  Gly Ala Thr Tyr Trp  Arg Glu Asn
    1925                 1930                 1935

Thr Met  Glu Phe Leu Glu Leu  Phe Leu Pro Glu Ser  Leu Thr Gly
    1940                 1945                 1950

Pro Gly  Ser Lys Ser Cys Asp  Gln His Tyr Gly Ser  Thr Cys Ala
    1955                 1960                 1965

Gly Gly  Ser Gln Arg Arg Val  Val Val Cys Gln Asp  Glu Asn Gly
    1970                 1975                 1980

Tyr Thr  Ala Asn Asp Cys Val  Glu Arg Ile Lys Pro  Asp Glu Gln
    1985                 1990                 1995

Arg Ala  Cys Glu Ser Gly Pro  Cys Pro Gln Trp Ala  Tyr Gly Asn
    2000                 2005                 2010

Trp Gly  Glu Cys Thr Lys Leu  Cys Gly Gly Gly Ile  Arg Thr Arg
    2015                 2020                 2025

Leu Val  Val Cys Gln Arg Ser  Asn Gly Glu Arg Phe  Pro Asp Leu
    2030                 2035                 2040

Ser Cys  Glu Ile Leu Asp Lys  Pro Pro Asp Arg Glu  Gln Cys Asn
    2045                 2050                 2055

Thr His  Ala Cys Pro His Asp  Ala Ala Trp Ser Thr  Gly Pro Trp
    2060                 2065                 2070

Ser Ser  Ser Met Trp Gln Val  Asn Asn Lys Thr Val  Thr Leu Gly
    2075                 2080                 2085

Asn Leu  Cys Ser Val Ser Cys  Gly Arg Gly His Lys  Gln Arg Asn
    2090                 2095                 2100

Val Tyr  Cys Met Ala Lys Asp  Gly Ser His Leu Glu  Ser Asp Tyr
    2105                 2110                 2115

Cys Lys  His Leu Ala Lys Pro  His Gly His Arg Lys  Cys Arg Gly
    2120                 2125                 2130

Gly Arg  Cys Pro Lys Trp Lys  Ala Gly Ala Trp Ser  Gln Lys Thr
    2135                 2140                 2145
```

```
Thr Asn  Ser Asp Cys Thr Glu  Ala Asp Cys Gly His  Leu Ala Glu
    2150                 2155                 2160

Ile Glu  Ser Gln Phe Ile Leu  Glu Val Leu Glu Glu  Arg Ala Val
    2165                 2170                 2175

Asp Glu  Ser Ser Arg Lys Tyr  Leu Cys Pro Phe Ala  Cys Leu Gln
    2180                 2185                 2190

Lys Cys  Ser Val Ser Cys Gly  Arg Gly Val Gln Gln  Arg His Val
    2195                 2200                 2205

Gly Cys  Gln Ile Gly Thr His  Lys Ile Ala Arg Glu  Thr Glu Cys
    2210                 2215                 2220

Asn Pro  Tyr Thr Arg Pro Glu  Ser Glu Arg Asp Cys  Gln Gly Pro
    2225                 2230                 2235

Arg Cys  Pro Leu Tyr Thr Trp  Arg Ala Glu Glu Trp  Gln Glu Thr
    2240                 2245                 2250

Tyr His  Gly Leu Leu Ser Pro  Ser Pro Ser Leu Cys  His Ala Lys
    2255                 2260                 2265

Leu Asn  Pro Ala Pro Arg Ser  Gly Lys Pro Gln Pro  Arg Cys His
    2270                 2275                 2280

Phe Leu  Ser Glu Ala Phe Ala  Asn His Thr Thr Pro  Leu Asn Leu
    2285                 2290                 2295

Ser Gln  Met Leu Leu His Ser  Ala Leu Thr Thr His  Ala Asp Tyr
    2300                 2305                 2310

Cys Thr  Leu Ala Val Asn Thr  Trp Asn Ser His Cys  Leu Phe Phe
    2315                 2320                 2325

Ser Ser  Met Leu Ser Val Ile
    2330                 2335
```

<210> 60
<211> 1072
<212> PRT
<213> Homo Sapiens

<400> 60

```
Met Gln Phe Val Ser Trp Ala Thr Leu Leu Thr Leu Leu Val Arg Asp
1               5                  10                  15
```

```
Leu Ala Glu Met Gly Ser Pro Asp Ala Ala Ala Ala Val Arg Lys Asp
            20                  25                  30

Arg Leu His Pro Arg Gln Val Lys Leu Leu Glu Thr Leu Gly Glu Tyr
        35                  40                  45

Glu Ile Val Ser Pro Ile Arg Val Asn Ala Leu Gly Glu Pro Phe Pro
    50                  55                  60

Thr Asn Val His Phe Lys Arg Thr Arg Arg Ser Ile Asn Ser Ala Thr
65                  70                  75                  80

Asp Pro Trp Pro Ala Phe Ala Ser Ser Ser Ser Ser Ser Thr Ser Ser
                85                  90                  95

Gln Ala His Tyr Arg Leu Ser Ala Phe Gly Gln Gln Phe Leu Phe Asn
            100                 105                 110

Leu Thr Ala Asn Ala Gly Phe Ile Ala Pro Leu Phe Thr Val Thr Leu
        115                 120                 125

Leu Gly Thr Pro Gly Val Asn Gln Thr Lys Phe Tyr Ser Glu Glu Glu
    130                 135                 140

Ala Glu Leu Lys His Cys Phe Tyr Lys Gly Tyr Val Asn Thr Asn Ser
145                 150                 155                 160

Glu His Thr Ala Val Ile Ser Leu Cys Ser Gly Met Leu Gly Thr Phe
                165                 170                 175

Arg Ser His Asp Gly Asp Tyr Phe Ile Glu Pro Leu Gln Ser Met Asp
            180                 185                 190

Glu Gln Glu Asp Glu Glu Glu Gln Asn Lys Pro His Ile Ile Tyr Arg
        195                 200                 205

Arg Ser Ala Pro Gln Arg Glu Pro Ser Thr Gly Arg His Ala Cys Asp
    210                 215                 220

Thr Ser Glu His Lys Asn Arg His Ser Lys Asp Lys Lys Lys Thr Arg
225                 230                 235                 240

Ala Arg Lys Trp Gly Glu Arg Ile Asn Leu Ala Gly Asp Val Ala Ala
                245                 250                 255
```

```
Leu Asn Ser Gly Leu Ala Thr Glu Ala Phe Ser Ala Tyr Gly Asn Lys
            260                 265                 270

Thr Asp Asn Thr Arg Glu Lys Arg Thr His Arg Arg Thr Lys Arg Phe
        275                 280                 285

Leu Ser Tyr Pro Arg Phe Val Glu Val Leu Val Val Ala Asp Asn Arg
    290                 295                 300

Met Val Ser Tyr His Gly Glu Asn Leu Gln His Tyr Ile Leu Thr Leu
305                 310                 315                 320

Met Ser Ile Val Ala Ser Ile Tyr Lys Asp Pro Ser Ile Gly Asn Leu
                325                 330                 335

Ile Asn Ile Val Ile Val Asn Leu Ile Val Ile His Asn Glu Gln Asp
            340                 345                 350

Gly Pro Ser Ile Ser Phe Asn Ala Gln Thr Thr Leu Lys Asn Leu Cys
        355                 360                 365

Gln Trp Gln His Ser Lys Asn Ser Pro Gly Gly Ile His His Asp Thr
    370                 375                 380

Ala Val Leu Leu Thr Arg Gln Asp Ile Cys Arg Ala His Asp Lys Cys
385                 390                 395                 400

Asp Thr Leu Gly Leu Ala Glu Leu Gly Thr Ile Cys Asp Pro Tyr Arg
                405                 410                 415

Ser Cys Ser Ile Ser Glu Asp Ser Gly Leu Ser Thr Ala Phe Thr Ile
            420                 425                 430

Ala His Glu Leu Gly His Val Phe Asn Met Pro His Asp Asp Asn Asn
        435                 440                 445

Lys Cys Lys Glu Glu Gly Val Lys Ser Pro Gln His Val Met Ala Pro
    450                 455                 460

Thr Leu Asn Phe Tyr Thr Asn Pro Trp Met Trp Ser Lys Cys Ser Arg
465                 470                 475                 480

Lys Tyr Ile Thr Glu Phe Leu Asp Thr Gly Tyr Gly Glu Cys Leu Leu
                485                 490                 495
```

Asn Glu Pro Glu Ser Arg Pro Tyr Pro Leu Pro Val Gln Leu Pro Gly
500 505 510

Ile Leu Tyr Asn Val Asn Lys Gln Cys Glu Leu Ile Phe Gly Pro Gly
515 520 525

Ser Gln Val Cys Pro Tyr Met Met Gln Cys Arg Arg Leu Trp Cys Asn
530 535 540

Asn Val Asn Gly Val His Lys Gly Cys Arg Thr Gln His Thr Pro Trp
545 550 555 560

Ala Asp Gly Thr Glu Cys Glu Pro Gly Lys His Cys Lys Tyr Gly Phe
565 570 575

Cys Val Pro Lys Glu Met Asp Val Pro Val Thr Asp Gly Ser Trp Gly
580 585 590

Ser Trp Ser Pro Phe Gly Thr Cys Ser Arg Thr Cys Gly Gly Gly Ile
595 600 605

Lys Thr Ala Ile Arg Glu Cys Asn Arg Pro Glu Pro Lys Asn Gly Gly
610 615 620

Lys Tyr Cys Val Gly Arg Arg Met Lys Phe Lys Ser Cys Asn Thr Glu
625 630 635 640

Pro Cys Leu Lys Gln Lys Arg Asp Phe Arg Asp Glu Gln Cys Ala His
645 650 655

Phe Asp Gly Lys His Phe Asn Ile Asn Gly Leu Leu Pro Asn Val Arg
660 665 670

Trp Val Pro Lys Tyr Ser Gly Ile Leu Met Lys Asp Arg Cys Lys Leu
675 680 685

Phe Cys Arg Val Ala Gly Asn Thr Ala Tyr Tyr Gln Leu Arg Asp Arg
690 695 700

Val Ile Asp Gly Thr Pro Cys Gly Gln Asp Thr Asn Asp Ile Cys Val
705 710 715 720

Gln Gly Leu Cys Arg Gln Ala Gly Cys Asp His Val Leu Asn Ser Lys
725 730 735

Ala Arg Arg Asp Lys Cys Gly Val Cys Gly Gly Asp Asn Ser Ser Cys

740 745 750

Lys Thr Val Ala Gly Thr Phe Asn Thr Val His Tyr Gly Tyr Asn Thr
755 760 765

Val Val Arg Ile Pro Ala Gly Ala Thr Asn Ile Asp Val Arg Gln His
770 775 780

Ser Phe Ser Gly Glu Thr Asp Asp Asp Asn Tyr Leu Ala Leu Ser Ser
785 790 795 800

Ser Lys Gly Glu Phe Leu Leu Asn Gly Asn Phe Val Val Thr Met Ala
805 810 815

Lys Arg Glu Ile Arg Ile Gly Asn Ala Val Val Glu Tyr Ser Gly Ser
820 825 830

Glu Thr Ala Val Glu Arg Ile Asn Ser Thr Asp Arg Ile Glu Gln Glu
835 840 845

Leu Leu Leu Gln Val Leu Ser Val Gly Lys Leu Tyr Asn Pro Asp Val
850 855 860

Arg Tyr Ser Phe Asn Ile Pro Ile Glu Asp Lys Pro Gln Gln Phe Tyr
865 870 875 880

Trp Asn Ser His Gly Pro Trp Gln Ala Cys Ser Lys Pro Cys Gln Gly
885 890 895

Glu Arg Lys Arg Lys Leu Val Cys Thr Arg Glu Ser Asp Gln Leu Thr
900 905 910

Val Ser Asp Gln Arg Cys Asp Arg Leu Pro Gln Pro Gly His Ile Thr
915 920 925

Glu Pro Cys Gly Thr Asp Cys Asp Leu Arg Trp His Val Ala Ser Arg
930 935 940

Ser Glu Cys Ser Ala Gln Cys Gly Leu Gly Tyr Arg Thr Leu Asp Ile
945 950 955 960

Tyr Cys Ala Lys Tyr Ser Arg Leu Asp Gly Lys Thr Glu Lys Val Asp
965 970 975

Asp Gly Phe Cys Ser Ser His Pro Lys Pro Ser Asn Arg Glu Lys Cys
980 985 990

Ser Gly Glu Cys Asn Thr Gly Gly Trp Arg Tyr Ser Ala Trp Thr Glu
995 1000 1005

Cys Ser Lys Ser Cys Asp Gly Gly Thr Gln Arg Arg Arg Ala Ile
1010 1015 1020

Cys Val Asn Thr Arg Asn Asp Val Leu Asp Asp Ser Lys Cys Thr
1025 1030 1035

His Gln Glu Lys Val Thr Ile Gln Arg Cys Ser Glu Phe Pro Cys
1040 1045 1050

Pro Gln Trp Lys Ser Gly Asp Trp Ser Glu Val Arg Trp Glu Gly
1055 1060 1065

Cys Tyr Phe Pro
1070

<210> 61
<211> 1356
<212> PRT
<213> Homo Sapiens

<400> 61

Met Gln Ser Lys Val Leu Leu Ala Val Ala Leu Trp Leu Cys Val Glu
1 5 10 15

Thr Arg Ala Ala Ser Val Gly Leu Pro Ser Val Ser Leu Asp Leu Pro
20 25 30

Arg Leu Ser Ile Gln Lys Asp Ile Leu Thr Ile Lys Ala Asn Thr Thr
35 40 45

Leu Gln Ile Thr Cys Arg Gly Gln Arg Asp Leu Asp Trp Leu Trp Pro
50 55 60

Asn Asn Gln Ser Gly Ser Glu Gln Arg Val Glu Val Thr Glu Cys Ser
65 70 75 80

Asp Gly Leu Phe Cys Lys Thr Leu Thr Ile Pro Lys Val Ile Gly Asn
85 90 95

Asp Thr Gly Ala Tyr Lys Cys Phe Tyr Arg Glu Thr Asp Leu Ala Ser
100 105 110

```
Val Ile Tyr Val Tyr Val Gln Asp Tyr Arg Ser Pro Phe Ile Ala Ser
        115                 120                 125

Val Ser Asp Gln His Gly Val Val Tyr Ile Thr Glu Asn Lys Asn Lys
    130                 135                 140

Thr Val Val Ile Pro Cys Leu Gly Ser Ile Ser Asn Leu Asn Val Ser
145                 150                 155                 160

Leu Cys Ala Arg Tyr Pro Glu Lys Arg Phe Val Pro Asp Gly Asn Arg
                165                 170                 175

Ile Ser Trp Asp Ser Lys Lys Gly Phe Thr Ile Pro Ser Tyr Met Ile
            180                 185                 190

Ser Tyr Ala Gly Met Val Phe Cys Glu Ala Lys Ile Asn Asp Glu Ser
        195                 200                 205

Tyr Gln Ser Ile Met Tyr Ile Val Val Val Val Gly Tyr Arg Ile Tyr
    210                 215                 220

Asp Val Val Leu Ser Pro Ser His Gly Ile Glu Leu Ser Val Gly Glu
225                 230                 235                 240

Lys Leu Val Leu Asn Cys Thr Ala Arg Thr Glu Leu Asn Val Gly Ile
                245                 250                 255

Asp Phe Asn Trp Glu Tyr Pro Ser Ser Lys His Gln His Lys Lys Leu
            260                 265                 270

Val Asn Arg Asp Leu Lys Thr Gln Ser Gly Ser Glu Met Lys Lys Phe
        275                 280                 285

Leu Ser Thr Leu Thr Ile Asp Gly Val Thr Arg Ser Asp Gln Gly Leu
    290                 295                 300

Tyr Thr Cys Ala Ala Ser Ser Gly Leu Met Thr Lys Lys Asn Ser Thr
305                 310                 315                 320

Phe Val Arg Val His Glu Lys Pro Phe Val Ala Phe Gly Ser Gly Met
                325                 330                 335

Glu Ser Leu Val Glu Ala Thr Val Gly Glu Arg Val Arg Ile Pro Ala
            340                 345                 350

Lys Tyr Leu Gly Tyr Pro Pro Pro Glu Ile Lys Trp Tyr Lys Asn Gly
```

355 360 365

Ile Pro Leu Glu Ser Asn His Thr Ile Lys Ala Gly His Val Leu Thr
370 375 380

Ile Met Glu Val Ser Glu Arg Asp Thr Gly Asn Tyr Thr Val Ile Leu
385 390 395 400

Thr Asn Pro Ile Ser Lys Glu Lys Gln Ser His Val Val Ser Leu Val
405 410 415

Val Tyr Val Pro Pro Gln Ile Gly Glu Lys Ser Leu Ile Ser Pro Val
420 425 430

Asp Ser Tyr Gln Tyr Gly Thr Thr Gln Thr Leu Thr Cys Thr Val Tyr
435 440 445

Ala Ile Pro Pro Pro His His Ile His Trp Tyr Trp Gln Leu Glu Glu
450 455 460

Glu Cys Ala Asn Glu Pro Ser Gln Ala Val Ser Val Thr Asn Pro Tyr
465 470 475 480

Pro Cys Glu Glu Trp Arg Ser Val Glu Asp Phe Gln Gly Gly Asn Lys
485 490 495

Ile Glu Val Asn Lys Asn Gln Phe Ala Leu Ile Glu Gly Lys Asn Lys
500 505 510

Thr Val Ser Thr Leu Val Ile Gln Ala Ala Asn Val Ser Ala Leu Tyr
515 520 525

Lys Cys Glu Ala Val Asn Lys Val Gly Arg Gly Glu Arg Val Ile Ser
530 535 540

Phe His Val Thr Arg Gly Pro Glu Ile Thr Leu Gln Pro Asp Met Gln
545 550 555 560

Pro Thr Glu Gln Glu Ser Val Ser Leu Trp Cys Thr Ala Asp Arg Ser
565 570 575

Thr Phe Glu Asn Leu Thr Trp Tyr Lys Leu Gly Pro Gln Pro Leu Pro
580 585 590

Ile His Val Gly Glu Leu Pro Thr Pro Val Cys Lys Asn Leu Asp Thr
595 600 605

```
Leu Trp Lys Leu Asn Ala Thr Met Phe Ser Asn Ser Thr Asn Asp Ile
    610                 615                 620

Leu Ile Met Glu Leu Lys Asn Ala Ser Leu Gln Asp Gln Gly Asp Tyr
625                 630                 635                 640

Val Cys Leu Ala Gln Asp Arg Lys Thr Lys Lys Arg His Cys Val Val
                645                 650                 655

Arg Gln Leu Thr Val Leu Glu Arg Val Ala Pro Thr Ile Thr Gly Asn
            660                 665                 670

Leu Glu Asn Gln Thr Thr Ser Ile Gly Glu Ser Ile Glu Val Ser Cys
        675                 680                 685

Thr Ala Ser Gly Asn Pro Pro Pro Gln Ile Met Trp Phe Lys Asp Asn
    690                 695                 700

Glu Thr Leu Val Glu Asp Ser Gly Ile Val Leu Lys Asp Gly Asn Arg
705                 710                 715                 720

Asn Leu Thr Ile Arg Arg Val Arg Lys Glu Asp Glu Gly Leu Tyr Thr
                725                 730                 735

Cys Gln Ala Cys Ser Val Leu Gly Cys Ala Lys Val Glu Ala Phe Phe
            740                 745                 750

Ile Ile Glu Gly Ala Gln Glu Lys Thr Asn Leu Glu Ile Ile Ile Leu
        755                 760                 765

Val Gly Thr Ala Val Ile Ala Met Phe Phe Trp Leu Leu Leu Val Ile
    770                 775                 780

Ile Leu Arg Thr Val Lys Arg Ala Asn Gly Gly Glu Leu Lys Thr Gly
785                 790                 795                 800

Tyr Leu Ser Ile Val Met Asp Pro Asp Glu Leu Pro Leu Asp Glu His
                805                 810                 815

Cys Glu Arg Leu Pro Tyr Asp Ala Ser Lys Trp Glu Phe Pro Arg Asp
            820                 825                 830

Arg Leu Lys Leu Gly Lys Pro Leu Gly Arg Gly Ala Phe Gly Gln Val
        835                 840                 845
```

Ile Glu Ala Asp Ala Phe Gly Ile Asp Lys Thr Ala Thr Cys Arg Thr
850 855 860

Val Ala Val Lys Met Leu Lys Glu Gly Ala Thr His Ser Glu His Arg
865 870 875 880

Ala Leu Met Ser Glu Leu Lys Ile Leu Ile His Ile Gly His His Leu
885 890 895

Asn Val Val Asn Leu Leu Gly Ala Cys Thr Lys Pro Gly Gly Pro Leu
900 905 910

Met Val Ile Val Glu Phe Cys Lys Phe Gly Asn Leu Ser Thr Tyr Leu
915 920 925

Arg Ser Lys Arg Asn Glu Phe Val Pro Tyr Lys Thr Lys Gly Ala Arg
930 935 940

Phe Arg Gln Gly Lys Asp Tyr Val Gly Ala Ile Pro Val Asp Leu Lys
945 950 955 960

Arg Arg Leu Asp Ser Ile Thr Ser Ser Gln Ser Ser Ala Ser Ser Gly
965 970 975

Phe Val Glu Glu Lys Ser Leu Ser Asp Val Glu Glu Glu Glu Ala Pro
980 985 990

Glu Asp Leu Tyr Lys Asp Phe Leu Thr Leu Glu His Leu Ile Cys Tyr
995 1000 1005

Ser Phe Gln Val Ala Lys Gly Met Glu Phe Leu Ala Ser Arg Lys
1010 1015 1020

Cys Ile His Arg Asp Leu Ala Ala Arg Asn Ile Leu Leu Ser Glu
1025 1030 1035

Lys Asn Val Val Lys Ile Cys Asp Phe Gly Leu Ala Arg Asp Ile
1040 1045 1050

Tyr Lys Asp Pro Asp Tyr Val Arg Lys Gly Asp Ala Arg Leu Pro
1055 1060 1065

Leu Lys Trp Met Ala Pro Glu Thr Ile Phe Asp Arg Val Tyr Thr
1070 1075 1080

Ile Gln Ser Asp Val Trp Ser Phe Gly Val Leu Leu Trp Glu Ile
1085 1090 1095

Phe Ser Leu Gly Ala Ser Pro Tyr Pro Gly Val Lys Ile Asp Glu
1100 1105 1110

Glu Phe Cys Arg Arg Leu Lys Glu Gly Thr Arg Met Arg Ala Pro
1115 1120 1125

Asp Tyr Thr Thr Pro Glu Met Tyr Gln Thr Met Leu Asp Cys Trp
1130 1135 1140

His Gly Glu Pro Ser Gln Arg Pro Thr Phe Ser Glu Leu Val Glu
1145 1150 1155

His Leu Gly Asn Leu Leu Gln Ala Asn Ala Gln Gln Asp Gly Lys
1160 1165 1170

Asp Tyr Ile Val Leu Pro Ile Ser Glu Thr Leu Ser Met Glu Glu
1175 1180 1185

Asp Ser Gly Leu Ser Leu Pro Thr Ser Pro Val Ser Cys Met Glu
1190 1195 1200

Glu Glu Glu Val Cys Asp Pro Lys Phe His Tyr Asp Asn Thr Ala
1205 1210 1215

Gly Ile Ser Gln Tyr Leu Gln Asn Ser Lys Arg Lys Ser Arg Pro
1220 1225 1230

Val Ser Val Lys Thr Phe Glu Asp Ile Pro Leu Glu Glu Pro Glu
1235 1240 1245

Val Lys Val Ile Pro Asp Asp Asn Gln Thr Asp Ser Gly Met Val
1250 1255 1260

Leu Ala Ser Glu Glu Leu Lys Thr Leu Glu Asp Arg Thr Lys Leu
1265 1270 1275

Ser Pro Ser Phe Gly Gly Met Val Pro Ser Lys Ser Arg Glu Ser
1280 1285 1290

Val Ala Ser Glu Gly Ser Asn Gln Thr Ser Gly Tyr Gln Ser Gly
1295 1300 1305

Tyr His Ser Asp Asp Thr Asp Thr Thr Val Tyr Ser Ser Glu Glu

1310 1315 1320

Ala Glu Leu Leu Lys Leu Ile Glu Ile Gly Val Gln Thr Gly Ser
1325 1330 1335

Thr Ala Gln Ile Leu Gln Pro Asp Ser Gly Thr Thr Leu Ser Ser
1340 1345 1350

Pro Pro Val
1355

<210> 62
<211> 468
<212> PRT
<213> Homo Sapiens

<400> 62

Met Gly Arg Gly Trp Gly Phe Leu Phe Gly Leu Leu Gly Ala Val Trp
1 5 10 15

Leu Leu Ser Ser Gly His Gly Glu Glu Gln Pro Pro Glu Thr Ala Ala
20 25 30

Gln Arg Cys Phe Cys Gln Val Ser Gly Tyr Leu Asp Asp Cys Thr Cys
35 40 45

Asp Val Glu Thr Ile Asp Arg Phe Asn Asn Tyr Arg Leu Phe Pro Arg
50 55 60

Leu Gln Lys Leu Leu Glu Ser Asp Tyr Phe Arg Tyr Tyr Lys Val Asn
65 70 75 80

Leu Lys Arg Pro Cys Pro Phe Trp Asn Asp Ile Ser Gln Cys Gly Arg
85 90 95

Arg Asp Cys Ala Val Lys Pro Cys Gln Ser Asp Glu Val Pro Asp Gly
100 105 110

Ile Lys Ser Ala Ser Tyr Lys Tyr Ser Glu Glu Ala Asn Asn Leu Ile
115 120 125

Glu Glu Cys Glu Gln Ala Glu Arg Leu Gly Ala Val Asp Glu Ser Leu
130 135 140

Ser Glu Glu Thr Gln Lys Ala Val Leu Gln Trp Thr Lys His Asp Asp
145 150 155 160

```
Ser Ser Asp Asn Phe Cys Glu Ala Asp Asp Ile Gln Ser Pro Glu Ala
                165                 170                 175

Glu Tyr Val Asp Leu Leu Leu Asn Pro Glu Arg Tyr Thr Gly Tyr Lys
            180                 185                 190

Gly Pro Asp Ala Trp Lys Ile Trp Asn Val Ile Tyr Glu Glu Asn Cys
        195                 200                 205

Phe Lys Pro Gln Thr Ile Lys Arg Pro Leu Asn Pro Leu Ala Ser Gly
    210                 215                 220

Gln Gly Thr Ser Glu Glu Asn Thr Phe Tyr Ser Trp Leu Glu Gly Leu
225                 230                 235                 240

Cys Val Glu Lys Arg Ala Phe Tyr Arg Leu Ile Ser Gly Leu His Ala
                245                 250                 255

Ser Ile Asn Val His Leu Ser Ala Arg Tyr Leu Leu Gln Glu Thr Trp
            260                 265                 270

Leu Glu Lys Lys Trp Gly His Asn Ile Thr Glu Phe Gln Gln Arg Phe
        275                 280                 285

Asp Gly Ile Leu Thr Glu Gly Glu Gly Pro Arg Arg Leu Lys Asn Leu
    290                 295                 300

Tyr Phe Leu Tyr Leu Ile Glu Leu Arg Ala Leu Ser Lys Val Leu Pro
305                 310                 315                 320

Phe Phe Glu Arg Pro Asp Phe Gln Leu Phe Thr Gly Asn Lys Ile Gln
                325                 330                 335

Asp Glu Glu Asn Lys Met Leu Leu Leu Glu Ile Leu His Glu Ile Lys
            340                 345                 350

Ser Phe Pro Leu His Phe Asp Glu Asn Ser Phe Phe Ala Gly Asp Lys
        355                 360                 365

Lys Glu Ala His Lys Leu Lys Glu Asp Phe Arg Leu His Phe Arg Asn
    370                 375                 380

Ile Ser Arg Ile Met Asp Cys Val Gly Cys Phe Lys Cys Arg Leu Trp
385                 390                 395                 400
```

Gly Lys Leu Gln Thr Gln Gly Leu Gly Thr Ala Leu Lys Ile Leu Phe
405 410 415

Ser Glu Lys Leu Ile Ala Asn Met Pro Glu Ser Gly Pro Ser Tyr Glu
420 425 430

Phe His Leu Thr Arg Gln Glu Ile Val Ser Leu Phe Asn Ala Phe Gly
435 440 445

Arg Ile Ser Thr Ser Val Lys Glu Leu Glu Asn Phe Arg Asn Leu Leu
450 455 460

Gln Asn Ile His
465

<210> 63
<211> 228
<212> PRT
<213> Homo Sapiens

<400> 63

Met Gln Pro Arg Arg Gln Arg Leu Pro Ala Pro Trp Ser Gly Pro Arg
1 5 10 15

Gly Pro Arg Pro Thr Ala Pro Leu Leu Ala Leu Leu Leu Leu Leu Ala
20 25 30

Pro Val Ala Ala Pro Ala Gly Ser Gly Gly Pro Asp Asp Pro Gly Gln
35 40 45

Pro Gln Asp Ala Gly Val Pro Arg Arg Leu Leu Gln Gln Lys Ala Arg
50 55 60

Ala Ala Leu His Phe Phe Asn Phe Arg Ser Gly Ser Pro Ser Ala Leu
65 70 75 80

Arg Val Leu Ala Glu Val Gln Glu Gly Arg Ala Trp Ile Asn Pro Lys
85 90 95

Glu Gly Cys Lys Val His Val Val Phe Ser Thr Glu Arg Tyr Asn Pro
100 105 110

Glu Ser Leu Leu Gln Glu Gly Glu Gly Arg Leu Gly Lys Cys Ser Ala
115 120 125

Arg Val Phe Phe Lys Asn Gln Lys Pro Arg Pro Thr Ile Asn Val Thr
130 135 140

Cys Thr Arg Leu Ile Glu Lys Lys Lys Arg Gln Gln Glu Asp Tyr Leu
145 150 155 160

Leu Tyr Lys Gln Met Lys Gln Leu Lys Asn Pro Leu Glu Ile Val Ser
165 170 175

Ile Pro Asp Asn His Gly His Ile Asp Pro Ser Leu Arg Leu Ile Trp
180 185 190

Asp Leu Ala Phe Leu Gly Ser Ser Tyr Val Met Trp Glu Met Thr Thr
195 200 205

Gln Val Ser His Tyr Tyr Leu Ala Gln Leu Thr Ser Val Arg Gln Trp
210 215 220

Val Arg Lys Thr
225

<210> 64
<211> 747
<212> PRT
<213> Homo Sapiens

<400> 64

Met Arg Arg Cys Asn Ser Gly Ser Gly Pro Pro Pro Ser Leu Leu Leu
1 5 10 15

Leu Leu Leu Trp Leu Leu Ala Val Pro Gly Ala Asn Ala Ala Pro Arg
20 25 30

Ser Ala Leu Tyr Ser Pro Ser Asp Pro Leu Thr Leu Leu Gln Ala Asp
35 40 45

Thr Val Arg Gly Ala Val Leu Gly Ser Arg Ser Ala Trp Ala Val Glu
50 55 60

Phe Phe Ala Ser Trp Cys Gly His Cys Ile Ala Phe Ala Pro Thr Trp
65 70 75 80

Lys Ala Leu Ala Glu Asp Val Lys Ala Trp Arg Pro Ala Leu Tyr Leu
85 90 95

Ala Ala Leu Asp Cys Ala Glu Glu Thr Asn Ser Ala Val Cys Arg Asp
100 105 110

```
Phe Asn Ile Pro Gly Phe Pro Thr Val Arg Phe Phe Lys Ala Phe Thr
        115                 120                 125

Lys Asn Gly Ser Gly Ala Val Phe Pro Val Ala Gly Ala Asp Val Gln
    130                 135                 140

Thr Leu Arg Glu Arg Leu Ile Asp Ala Leu Glu Ser His His Asp Thr
145                 150                 155                 160

Trp Pro Pro Ala Cys Pro Pro Leu Glu Pro Ala Lys Leu Glu Glu Ile
                165                 170                 175

Asp Gly Phe Phe Ala Arg Asn Asn Glu Glu Tyr Leu Ala Leu Ile Phe
            180                 185                 190

Glu Lys Gly Gly Ser Tyr Leu Gly Arg Glu Val Ala Leu Asp Leu Ser
        195                 200                 205

Gln His Lys Gly Val Ala Val Arg Arg Val Leu Asn Thr Glu Ala Asn
    210                 215                 220

Val Val Arg Lys Phe Gly Val Thr Asp Phe Pro Ser Cys Tyr Leu Leu
225                 230                 235                 240

Phe Arg Asn Gly Ser Val Ser Arg Val Pro Val Leu Met Glu Ser Arg
                245                 250                 255

Ser Phe Tyr Thr Ala Tyr Leu Gln Arg Leu Ser Gly Leu Thr Arg Glu
            260                 265                 270

Ala Ala Gln Thr Thr Val Ala Pro Thr Thr Ala Asn Lys Ile Ala Pro
        275                 280                 285

Thr Val Trp Lys Leu Ala Asp Arg Ser Lys Ile Tyr Met Ala Asp Leu
    290                 295                 300

Glu Ser Ala Leu His Tyr Ile Leu Arg Ile Glu Val Gly Arg Phe Pro
305                 310                 315                 320

Val Leu Glu Gly Gln Arg Leu Val Ala Leu Lys Lys Phe Val Ala Val
                325                 330                 335

Leu Ala Lys Tyr Phe Pro Gly Arg Pro Leu Val Gln Asn Phe Leu His
            340                 345                 350

Ser Val Asn Glu Trp Leu Lys Arg Gln Lys Arg Asn Lys Ile Pro Tyr
```

355 360 365

Ser Phe Phe Lys Thr Ala Leu Asp Asp Arg Lys Glu Gly Ala Val Leu
370 375 380

Ala Lys Lys Val Asn Trp Ile Gly Cys Gln Gly Ser Glu Pro His Phe
385 390 395 400

Arg Gly Phe Pro Cys Ser Leu Trp Val Leu Phe His Phe Leu Thr Val
405 410 415

Gln Ala Ala Arg Gln Asn Val Asp His Ser Gln Glu Ala Ala Lys Ala
420 425 430

Lys Glu Val Leu Pro Ala Ile Arg Gly Tyr Val His Tyr Phe Phe Gly
435 440 445

Cys Arg Asp Cys Ala Ser His Phe Glu Gln Met Ala Ala Ala Ser Met
450 455 460

His Arg Val Gly Ser Pro Asn Ala Ala Val Leu Trp Leu Trp Ser Ser
465 470 475 480

His Asn Arg Val Asn Ala Arg Leu Ala Gly Ala Pro Ser Glu Asp Pro
485 490 495

Gln Phe Pro Lys Val Gln Trp Pro Pro Arg Glu Leu Cys Ser Ala Cys
500 505 510

His Asn Glu Arg Leu Asp Val Pro Val Trp Asp Val Glu Ala Thr Leu
515 520 525

Asn Phe Leu Lys Ala His Phe Ser Pro Ser Asn Ile Ile Leu Asp Phe
530 535 540

Pro Ala Ala Gly Ser Ala Ala Arg Arg Asp Val Gln Asn Val Ala Ala
545 550 555 560

Ala Pro Glu Leu Ala Met Gly Ala Leu Glu Leu Glu Ser Arg Asn Ser
565 570 575

Thr Leu Asp Pro Gly Lys Pro Glu Met Met Lys Ser Pro Thr Asn Thr
580 585 590

Thr Pro His Val Pro Ala Glu Gly Pro Glu Ala Ser Arg Pro Pro Lys
595 600 605

Leu His Pro Gly Leu Arg Ala Ala Pro Gly Gln Glu Pro Pro Glu His
610 615 620

Met Ala Glu Leu Gln Arg Asn Glu Gln Glu Gln Pro Leu Gly Gln Trp
625 630 635 640

His Leu Ser Lys Arg Asp Thr Gly Ala Ala Leu Leu Ala Glu Ser Arg
645 650 655

Ala Glu Lys Asn Arg Leu Trp Gly Pro Leu Glu Val Arg Arg Val Gly
660 665 670

Arg Ser Ser Lys Gln Leu Val Asp Ile Pro Glu Gly Gln Leu Glu Ala
675 680 685

Arg Ala Gly Arg Gly Arg Gly Gln Trp Leu Gln Val Leu Gly Gly Gly
690 695 700

Phe Ser Tyr Leu Asp Ile Ser Leu Cys Val Gly Leu Tyr Ser Leu Ser
705 710 715 720

Phe Met Gly Leu Leu Ala Met Tyr Thr Tyr Phe Gln Ala Lys Ile Arg
725 730 735

Ala Leu Lys Gly His Ala Gly His Pro Ala Ala
740 745

<210> 65
<211> 1163
<212> PRT
<213> Homo Sapiens

<400> 65

Met Val Trp Cys Leu Gly Leu Ala Val Leu Ser Leu Val Ile Ser Gln
1 5 10 15

Gly Ala Asp Gly Arg Gly Lys Pro Glu Val Val Ser Val Val Gly Arg
20 25 30

Ala Glu Glu Ser Val Val Leu Gly Cys Asp Leu Leu Pro Pro Ala Gly
35 40 45

Arg Pro Pro Leu His Val Ile Glu Trp Leu Arg Phe Gly Phe Leu Leu
50 55 60

```
Pro Ile Phe Ile Gln Phe Gly Leu Tyr Ser Pro Arg Ile Asp Pro Asp
65                  70                  75                  80

Tyr Val Gly Arg Val Arg Leu Gln Lys Gly Ala Ser Leu Gln Ile Glu
                85                  90                  95

Gly Leu Arg Val Glu Asp Gln Gly Trp Tyr Glu Cys Arg Val Phe Phe
            100                 105                 110

Leu Asp Gln His Ile Pro Glu Asp Asp Phe Ala Asn Gly Ser Trp Val
        115                 120                 125

His Leu Thr Val Asn Ser Pro Pro Gln Phe Gln Glu Thr Pro Pro Ala
    130                 135                 140

Val Leu Glu Val Gln Glu Leu Glu Pro Val Thr Leu Arg Cys Val Ala
145                 150                 155                 160

Arg Gly Ser Pro Leu Pro His Val Thr Trp Lys Leu Arg Gly Lys Asp
                165                 170                 175

Leu Gly Gln Gly Gln Gly Gln Val Gln Val Gln Asn Gly Thr Leu Arg
            180                 185                 190

Ile Arg Arg Val Glu Arg Gly Ser Ser Gly Val Tyr Thr Cys Gln Ala
        195                 200                 205

Ser Ser Thr Glu Gly Ser Ala Thr His Ala Thr Gln Leu Leu Val Leu
    210                 215                 220

Gly Pro Pro Val Ile Val Val Pro Pro Lys Asn Ser Thr Val Asn Ala
225                 230                 235                 240

Ser Gln Asp Val Ser Leu Ala Cys His Ala Glu Ala Tyr Pro Ala Asn
                245                 250                 255

Leu Thr Tyr Ser Trp Phe Gln Asp Asn Ile Asn Val Phe His Ile Ser
            260                 265                 270

Arg Leu Gln Pro Arg Val Gln Ile Leu Val Asp Gly Ser Leu Arg Leu
        275                 280                 285

Leu Ala Thr Gln Pro Asp Asp Ala Gly Cys Tyr Thr Cys Val Pro Ser
    290                 295                 300

Asn Gly Leu Leu His Pro Pro Ser Ala Ser Ala Tyr Leu Thr Val Leu
```

```
305                 310                 315                 320

Cys Met Pro Gly Val Ile Arg Cys Pro Val Arg Ala Asn Pro Pro Leu
                325                 330                 335

Leu Phe Val Ser Trp Thr Lys Asp Gly Lys Ala Leu Gln Leu Asp Lys
            340                 345                 350

Phe Pro Gly Trp Ser Gln Gly Thr Glu Gly Ser Leu Ile Ile Ala Leu
        355                 360                 365

Gly Asn Glu Asp Ala Leu Gly Glu Tyr Ser Cys Thr Pro Tyr Asn Ser
    370                 375                 380

Leu Gly Thr Ala Gly Pro Ser Pro Val Thr Arg Val Leu Leu Lys Ala
385                 390                 395                 400

Pro Pro Ala Phe Ile Glu Arg Pro Lys Glu Glu Tyr Phe Gln Glu Val
                405                 410                 415

Gly Arg Glu Leu Leu Ile Pro Cys Ser Ala Gln Gly Asp Pro Pro Pro
            420                 425                 430

Val Val Ser Trp Thr Lys Val Gly Arg Gly Leu Gln Gly Gln Ala Gln
        435                 440                 445

Val Asp Ser Asn Ser Ser Leu Ile Leu Arg Pro Leu Thr Lys Glu Ala
    450                 455                 460

His Gly His Trp Glu Cys Ser Ala Ser Asn Ala Val Ala Arg Val Ala
465                 470                 475                 480

Thr Ser Thr Asn Val Tyr Val Leu Gly Thr Ser Pro His Val Val Thr
                485                 490                 495

Asn Val Ser Val Val Ala Leu Pro Lys Gly Ala Asn Val Ser Trp Glu
            500                 505                 510

Pro Gly Phe Asp Gly Gly Tyr Leu Gln Arg Phe Ser Val Trp Tyr Thr
        515                 520                 525

Pro Leu Ala Lys Arg Pro Asp Arg Met His His Asp Trp Val Ser Leu
    530                 535                 540

Ala Val Pro Val Gly Ala Ala His Leu Leu Val Pro Gly Leu Gln Pro
545                 550                 555                 560
```

His Thr Gln Tyr Gln Phe Ser Val Leu Ala Gln Asn Lys Leu Gly Ser
565 570 575

Gly Pro Phe Ser Glu Ile Val Leu Ser Ala Pro Glu Gly Leu Pro Thr
580 585 590

Thr Pro Ala Ala Pro Gly Leu Pro Pro Thr Glu Ile Pro Pro Pro Leu
595 600 605

Ser Pro Pro Arg Gly Leu Val Ala Val Arg Thr Pro Arg Gly Val Leu
610 615 620

Leu His Trp Asp Pro Pro Glu Leu Val Pro Lys Arg Leu Asp Gly Tyr
625 630 635 640

Val Leu Glu Gly Arg Gln Gly Ser Gln Gly Trp Glu Val Leu Asp Pro
645 650 655

Ala Val Ala Gly Thr Glu Thr Glu Leu Leu Val Pro Gly Leu Ile Lys
660 665 670

Asp Val Leu Tyr Glu Phe Arg Leu Val Ala Phe Ala Gly Ser Phe Val
675 680 685

Ser Asp Pro Ser Asn Thr Ala Asn Val Ser Thr Ser Gly Leu Glu Val
690 695 700

Tyr Pro Ser Arg Thr Gln Leu Pro Gly Leu Leu Pro Gln Pro Val Leu
705 710 715 720

Ala Gly Val Val Gly Gly Val Cys Phe Leu Gly Val Ala Val Leu Val
725 730 735

Ser Ile Leu Ala Gly Cys Leu Leu Asn Arg Arg Arg Ala Ala Arg Arg
740 745 750

Arg Arg Lys Arg Leu Arg Gln Asp Pro Pro Leu Ile Phe Ser Pro Thr
755 760 765

Gly Lys Ser Ala Ala Pro Ser Ala Leu Gly Ser Gly Ser Pro Asp Ser
770 775 780

Val Ala Lys Leu Lys Leu Gln Gly Ser Pro Val Pro Ser Leu Arg Gln
785 790 795 800

Ser Leu Leu Trp Gly Asp Pro Ala Gly Thr Pro Ser Pro His Pro Asp
805 810 815

Pro Pro Ser Ser Arg Gly Pro Leu Pro Leu Glu Pro Ile Cys Arg Gly
820 825 830

Pro Asp Gly Arg Phe Val Met Gly Pro Thr Val Ala Ala Pro Gln Glu
835 840 845

Arg Ser Gly Arg Glu Gln Ala Glu Pro Arg Thr Pro Ala Gln Arg Leu
850 855 860

Ala Arg Ser Phe Asp Cys Ser Ser Ser Ser Pro Ser Gly Ala Pro Gln
865 870 875 880

Pro Leu Cys Ile Glu Asp Ile Ser Pro Val Ala Pro Pro Pro Ala Ala
885 890 895

Pro Pro Ser Pro Leu Pro Gly Pro Gly Pro Leu Leu Gln Tyr Leu Ser
900 905 910

Leu Pro Phe Phe Arg Glu Met Asn Val Asp Gly Asp Trp Pro Pro Leu
915 920 925

Glu Glu Pro Ser Pro Ala Ala Pro Pro Asp Tyr Met Asp Thr Arg Arg
930 935 940

Cys Pro Thr Ser Ser Phe Leu Arg Ser Pro Glu Thr Pro Pro Val Ser
945 950 955 960

Pro Arg Glu Ser Leu Pro Gly Ala Val Val Gly Ala Gly Ala Thr Ala
965 970 975

Glu Pro Pro Tyr Thr Ala Leu Ala Asp Trp Thr Leu Arg Glu Arg Leu
980 985 990

Leu Pro Gly Leu Leu Pro Ala Ala Pro Arg Gly Ser Leu Thr Ser Gln
995 1000 1005

Ser Ser Gly Arg Gly Ser Ala Ser Phe Leu Arg Pro Pro Ser Thr
1010 1015 1020

Ala Pro Ser Ala Gly Gly Ser Tyr Leu Ser Pro Ala Pro Gly Asp
1025 1030 1035

Thr Ser Ser Trp Ala Ser Gly Pro Glu Arg Trp Pro Arg Arg Glu
1040 1045 1050

His Val Val Thr Val Ser Lys Arg Arg Asn Thr Ser Val Asp Glu
1055 1060 1065

Asn Tyr Glu Trp Asp Ser Glu Phe Pro Gly Asp Met Glu Leu Leu
1070 1075 1080

Glu Thr Leu His Leu Gly Leu Ala Ser Ser Arg Leu Arg Pro Glu
1085 1090 1095

Ala Glu Thr Glu Leu Gly Val Lys Thr Pro Glu Glu Gly Cys Leu
1100 1105 1110

Leu Asn Thr Ala His Val Thr Gly Pro Glu Ala Arg Cys Ala Ala
1115 1120 1125

Leu Arg Glu Glu Phe Leu Ala Phe Arg Arg Arg Arg Asp Ala Thr
1130 1135 1140

Arg Ala Arg Leu Pro Ala Tyr Arg Gln Pro Val Pro His Pro Glu
1145 1150 1155

Gln Ala Thr Leu Leu
1160

<210> 66
<211> 87
<212> PRT
<213> Homo Sapiens

<400> 66

Met Ala Gly Ala Ser Leu Gly Ala Arg Phe Tyr Arg Gln Ile Lys Arg
1 5 10 15

His Pro Gly Ile Ile Pro Met Ile Gly Leu Ile Cys Leu Gly Met Gly
20 25 30

Ser Ala Ala Leu Tyr Leu Leu Arg Leu Ala Leu Arg Ser Pro Asp Val
35 40 45

Cys Trp Asp Arg Lys Asn Asn Pro Glu Pro Trp Asn Arg Leu Ser Pro
50 55 60

Asn Asp Gln Tyr Lys Phe Leu Ala Val Ser Thr Asp Tyr Lys Lys Leu
65 70 75 80

Lys Lys Asp Arg Pro Asp Phe
85

<210> 67
<211> 1241
<212> PRT
<213> Homo Sapiens

<400> 67

Met Ile Met Phe Pro Leu Phe Gly Lys Ile Ser Leu Gly Ile Leu Ile
1 5 10 15

Phe Val Leu Ile Glu Gly Asp Phe Pro Ser Leu Thr Ala Gln Thr Tyr
20 25 30

Leu Ser Ile Glu Glu Ile Gln Glu Pro Lys Ser Ala Val Ser Phe Leu
35 40 45

Leu Pro Glu Glu Ser Thr Asp Leu Ser Leu Ala Thr Lys Lys Lys Gln
50 55 60

Pro Leu Asp Arg Arg Glu Thr Glu Arg Gln Trp Leu Ile Arg Arg Arg
65 70 75 80

Arg Ser Ile Leu Phe Pro Asn Gly Val Lys Ile Cys Pro Asp Glu Ser
85 90 95

Val Ala Glu Ala Val Ala Asn His Val Lys Tyr Phe Lys Val Arg Val
100 105 110

Cys Gln Glu Ala Val Trp Glu Ala Phe Arg Thr Phe Trp Asp Arg Leu
115 120 125

Pro Gly Arg Glu Glu Tyr His Tyr Trp Met Asn Leu Cys Glu Asp Gly
130 135 140

Val Thr Ser Ile Phe Glu Met Gly Thr Asn Phe Ser Glu Ser Val Glu
145 150 155 160

His Arg Ser Leu Ile Met Lys Lys Leu Thr Tyr Ala Lys Glu Thr Val
165 170 175

Ser Ser Ser Glu Leu Ser Ser Pro Val Pro Val Gly Asp Thr Ser Thr
180 185 190

Leu Gly Asp Thr Thr Leu Ser Val Pro His Pro Glu Val Asp Ala Tyr
195 200 205

Glu Gly Ala Ser Glu Ser Ser Leu Glu Arg Pro Glu Glu Ser Ile Ser
210 215 220

Asn Glu Ile Glu Asn Val Ile Glu Glu Ala Thr Lys Pro Ala Gly Glu
225 230 235 240

Gln Ile Ala Glu Phe Ser Ile His Leu Leu Gly Lys Gln Tyr Arg Glu
245 250 255

Glu Leu Gln Asp Ser Ser Ser Phe His His Gln His Leu Glu Glu Glu
260 265 270

Phe Ile Ser Glu Val Glu Asn Ala Phe Thr Gly Leu Pro Gly Tyr Lys
275 280 285

Glu Ile Arg Val Leu Glu Phe Arg Ser Pro Lys Glu Asn Asp Ser Gly
290 295 300

Val Asp Val Tyr Tyr Ala Val Thr Phe Asn Gly Glu Ala Ile Ser Asn
305 310 315 320

Thr Thr Trp Asp Leu Ile Ser Leu His Ser Asn Lys Val Glu Asn His
325 330 335

Gly Leu Val Glu Leu Asp Asp Lys Pro Thr Val Val Tyr Thr Ile Ser
340 345 350

Asn Phe Arg Asp Tyr Ile Ala Glu Thr Leu Gln Gln Asn Phe Leu Leu
355 360 365

Gly Asn Ser Ser Leu Asn Pro Asp Pro Asp Ser Leu Gln Leu Ile Asn
370 375 380

Val Arg Gly Val Leu Arg His Gln Thr Glu Asp Leu Val Trp Asn Thr
385 390 395 400

Gln Ser Ser Ser Leu Gln Ala Thr Pro Ser Ser Ile Leu Asp Asn Thr
405 410 415

Phe Gln Ala Ala Trp Pro Ser Ala Asp Glu Ser Ile Thr Ser Ser Ile
420 425 430

Pro Pro Leu Asp Phe Ser Ser Gly Pro Pro Ser Ala Thr Gly Arg Glu

435 440 445

Leu Trp Ser Glu Ser Pro Leu Gly Asp Leu Val Ser Thr His Lys Leu
450 455 460

Ala Phe Pro Ser Lys Met Gly Leu Ser Ser Ser Pro Glu Val Leu Glu
465 470 475 480

Val Ser Ser Leu Thr Leu His Ser Val Thr Pro Ala Val Leu Gln Thr
485 490 495

Gly Leu Pro Val Ala Ser Glu Glu Arg Thr Ser Gly Ser His Leu Val
500 505 510

Glu Asp Gly Leu Ala Asn Val Glu Glu Ser Glu Asp Phe Leu Ser Ile
515 520 525

Asp Ser Leu Pro Ser Ser Ser Phe Thr Gln Pro Val Pro Lys Glu Thr
530 535 540

Ile Pro Ser Met Glu Asp Ser Asp Val Ser Leu Thr Ser Ser Pro Tyr
545 550 555 560

Leu Thr Ser Ser Ile Pro Phe Gly Leu Asp Ser Leu Thr Ser Lys Val
565 570 575

Lys Asp Gln Leu Lys Val Ser Pro Phe Leu Pro Asp Ala Ser Met Glu
580 585 590

Lys Glu Leu Ile Phe Asp Gly Gly Leu Gly Ser Gly Ser Gly Gln Lys
595 600 605

Val Asp Leu Ile Thr Trp Pro Trp Ser Glu Thr Ser Ser Glu Lys Ser
610 615 620

Ala Glu Pro Leu Ser Lys Pro Trp Leu Glu Asp Asp Asp Ser Leu Leu
625 630 635 640

Pro Ala Glu Ile Glu Asp Lys Lys Leu Val Leu Val Asp Lys Met Asp
645 650 655

Ser Thr Asp Gln Ile Ser Lys His Ser Lys Tyr Glu His Asp Asp Arg
660 665 670

Ser Thr His Phe Pro Glu Glu Glu Pro Leu Ser Gly Pro Ala Val Pro
675 680 685

```
Ile Phe Ala Asp Thr Ala Ala Glu Ser Ala Ser Leu Thr Leu Pro Lys
    690                 695                 700

His Ile Ser Glu Val Pro Gly Val Asp Asp Cys Ser Val Thr Lys Ala
705                 710                 715                 720

Pro Leu Ile Leu Thr Ser Val Ala Ile Ser Ala Ser Thr Asp Lys Ser
                725                 730                 735

Asp Gln Ala Asp Ala Ile Leu Arg Glu Asp Met Glu Gln Ile Thr Glu
            740                 745                 750

Ser Ser Asn Tyr Glu Trp Phe Asp Ser Glu Val Ser Met Val Lys Pro
        755                 760                 765

Asp Met Gln Thr Leu Trp Thr Ile Leu Pro Glu Ser Glu Arg Val Trp
    770                 775                 780

Thr Arg Thr Ser Ser Leu Glu Lys Leu Ser Arg Asp Ile Leu Ala Ser
785                 790                 795                 800

Thr Pro Gln Ser Ala Asp Arg Leu Trp Leu Ser Val Thr Gln Ser Thr
                805                 810                 815

Lys Leu Pro Pro Thr Thr Ile Ser Thr Leu Leu Glu Asp Glu Val Ile
            820                 825                 830

Met Gly Val Gln Asp Ile Ser Leu Glu Leu Asp Arg Ile Gly Thr Asp
        835                 840                 845

Tyr Tyr Gln Pro Glu Gln Val Gln Glu Gln Asn Gly Lys Val Gly Ser
    850                 855                 860

Tyr Val Glu Met Ser Thr Ser Val His Ser Thr Glu Met Val Ser Val
865                 870                 875                 880

Ala Trp Pro Thr Glu Gly Gly Asp Asp Leu Ser Tyr Thr Gln Thr Ser
                885                 890                 895

Gly Ala Leu Val Val Phe Phe Ser Leu Arg Val Thr Asn Met Met Phe
            900                 905                 910

Ser Glu Asp Leu Phe Asn Lys Asn Ser Leu Glu Tyr Lys Ala Leu Glu
        915                 920                 925
```

```
Gln Arg Phe Leu Glu Leu Leu Val Pro Tyr Leu Gln Ser Asn Leu Thr
    930                 935                 940

Gly Phe Gln Asn Leu Glu Ile Leu Asn Phe Arg Asn Gly Ser Ile Val
945                 950                 955                 960

Val Asn Ser Arg Met Lys Phe Ala Asn Ser Val Pro Pro Asn Val Asn
                965                 970                 975

Asn Ala Val Tyr Met Ile Leu Glu Asp Phe Cys Thr Thr Ala Tyr Asn
            980                 985                 990

Thr Met Asn Leu Ala Ile Asp Lys  Tyr Ser Leu Asp Val  Glu Ser Gly
        995                 1000                1005

Asp Glu  Ala Asn Pro Cys Lys  Phe Gln Ala Cys Asn  Glu Phe Ser
    1010                1015                1020

Glu Cys  Leu Val Asn Pro Trp  Ser Gly Glu Ala Lys  Cys Arg Cys
    1025                1030                1035

Phe Pro  Gly Tyr Leu Ser Val  Glu Glu Arg Pro Cys  Gln Ser Leu
    1040                1045                1050

Cys Asp  Leu Gln Pro Asp Phe  Cys Leu Asn Asp Gly  Lys Cys Asp
    1055                1060                1065

Ile Met  Pro Gly His Gly Ala  Ile Cys Arg Cys Arg  Val Gly Glu
    1070                1075                1080

Asn Trp  Trp Tyr Arg Gly Lys  His Cys Glu Glu Phe  Val Ser Glu
    1085                1090                1095

Pro Val  Ile Ile Gly Ile Thr  Ile Ala Ser Val Val  Gly Leu Leu
    1100                1105                1110

Val Ile  Phe Ser Ala Ile Ile  Tyr Phe Phe Ile Arg  Thr Leu Gln
    1115                1120                1125

Ala His  His Asp Arg Ser Glu  Arg Glu Ser Pro Phe  Ser Gly Ser
    1130                1135                1140

Ser Arg  Gln Pro Asp Ser Leu  Ser Ser Ile Glu Asn  Ala Val Lys
    1145                1150                1155
```

Tyr Asn Pro Val Tyr Glu Ser His Arg Ala Gly Cys Glu Lys Tyr
1160 1165 1170

Glu Gly Pro Tyr Pro Gln His Pro Phe Tyr Ser Ser Ala Ser Gly
1175 1180 1185

Asp Val Ile Gly Gly Leu Ser Arg Glu Glu Ile Arg Gln Met Tyr
1190 1195 1200

Glu Ser Ser Glu Leu Ser Arg Glu Glu Ile Gln Glu Arg Met Arg
1205 1210 1215

Val Leu Glu Leu Tyr Ala Asn Asp Pro Glu Phe Ala Ala Phe Val
1220 1225 1230

Arg Glu Gln Gln Val Glu Glu Val
1235 1240

<210> 68
<211> 211
<212> PRT
<213> Homo Sapiens

<400> 68

Met Ala Asn Ala Gly Leu Gln Leu Leu Gly Phe Ile Leu Ala Phe Leu
1 5 10 15

Gly Trp Ile Gly Ala Ile Val Ser Thr Ala Leu Pro Gln Trp Arg Ile
20 25 30

Tyr Ser Tyr Ala Gly Asp Asn Ile Val Thr Ala Gln Ala Met Tyr Glu
35 40 45

Gly Leu Trp Met Ser Cys Val Ser Gln Ser Thr Gly Gln Ile Gln Cys
50 55 60

Lys Val Phe Asp Ser Leu Leu Asn Leu Ser Ser Thr Leu Gln Ala Thr
65 70 75 80

Arg Ala Leu Met Val Val Gly Ile Leu Leu Gly Val Ile Ala Ile Phe
85 90 95

Val Ala Thr Val Gly Met Lys Cys Met Lys Cys Leu Glu Asp Asp Glu
100 105 110

Val Gln Lys Met Arg Met Ala Val Ile Gly Gly Ala Ile Phe Leu Leu
115 120 125

Ala Gly Leu Ala Ile Leu Val Ala Thr Ala Trp Tyr Gly Asn Arg Ile
130 135 140

Val Gln Glu Phe Tyr Asp Pro Met Thr Pro Val Asn Ala Arg Tyr Glu
145 150 155 160

Phe Gly Gln Ala Leu Phe Thr Gly Trp Ala Ala Ala Ser Leu Cys Leu
165 170 175

Leu Gly Gly Ala Leu Leu Cys Cys Ser Cys Pro Arg Lys Thr Thr Ser
180 185 190

Tyr Pro Thr Pro Arg Pro Tyr Pro Lys Pro Ala Pro Ser Ser Gly Lys
195 200 205

Asp Tyr Val
210

<210> 69
<211> 360
<212> PRT
<213> Homo Sapiens

<400> 69

Met Asp Leu His Leu Phe Asp Tyr Ser Glu Pro Gly Asn Phe Ser Asp
1 5 10 15

Ile Ser Trp Pro Cys Asn Ser Ser Asp Cys Ile Val Val Asp Thr Val
20 25 30

Met Cys Pro Asn Met Pro Asn Lys Ser Val Leu Leu Tyr Thr Leu Ser
35 40 45

Phe Ile Tyr Ile Phe Ile Phe Val Ile Gly Met Ile Ala Asn Ser Val
50 55 60

Val Val Trp Val Asn Ile Gln Ala Lys Thr Thr Gly Tyr Asp Thr His
65 70 75 80

Cys Tyr Ile Leu Asn Leu Ala Ile Ala Asp Leu Trp Val Val Leu Thr
85 90 95

Ile Pro Val Trp Val Val Ser Leu Val Gln His Asn Gln Trp Pro Met
100 105 110

Gly Glu Leu Thr Cys Lys Val Thr His Leu Ile Phe Ser Ile Asn Leu
115 120 125

Phe Gly Ser Ile Phe Phe Leu Thr Cys Met Ser Val Asp Arg Tyr Leu
130 135 140

Ser Ile Thr Tyr Phe Thr Asn Thr Pro Ser Ser Arg Lys Lys Met Val
145 150 155 160

Arg Arg Val Val Cys Ile Leu Val Trp Leu Leu Ala Phe Cys Val Ser
165 170 175

Leu Pro Asp Thr Tyr Tyr Leu Lys Thr Val Thr Ser Ala Ser Asn Asn
180 185 190

Glu Thr Tyr Cys Arg Ser Phe Tyr Pro Glu His Ser Ile Lys Glu Trp
195 200 205

Leu Ile Gly Met Glu Leu Val Ser Val Val Leu Gly Phe Ala Val Pro
210 215 220

Phe Ser Ile Ile Ala Val Phe Tyr Phe Leu Leu Ala Arg Ala Ile Ser
225 230 235 240

Ala Ser Ser Asp Gln Glu Lys His Ser Ser Arg Lys Ile Ile Phe Ser
245 250 255

Tyr Val Val Val Phe Leu Val Cys Trp Leu Pro Tyr His Val Ala Val
260 265 270

Leu Leu Asp Ile Phe Ser Ile Leu His Tyr Ile Pro Phe Thr Cys Arg
275 280 285

Leu Glu His Ala Leu Phe Thr Ala Leu His Val Thr Gln Cys Leu Ser
290 295 300

Leu Val His Cys Cys Val Asn Pro Val Leu Tyr Ser Phe Ile Asn Arg
305 310 315 320

Asn Tyr Arg Tyr Glu Leu Met Lys Ala Phe Ile Phe Lys Tyr Ser Ala
325 330 335

Lys Thr Gly Leu Thr Lys Leu Ile Asp Ala Ser Arg Val Ser Glu Thr
340 345 350

Glu Tyr Ser Ala Leu Glu Gln Ser

355 360

<210> 70
<211> 2273
<212> PRT
<213> Homo Sapiens

<400> 70

Met Gly Phe Val Arg Gln Ile Gln Leu Leu Leu Trp Lys Asn Trp Thr
1 5 10 15

Leu Arg Lys Arg Gln Lys Ile Arg Phe Val Val Glu Leu Val Trp Pro
20 25 30

Leu Ser Leu Phe Leu Val Leu Ile Trp Leu Arg Asn Ala Asn Pro Leu
35 40 45

Tyr Ser His His Glu Cys His Phe Pro Asn Lys Ala Met Pro Ser Ala
50 55 60

Gly Met Leu Pro Trp Leu Gln Gly Ile Phe Cys Asn Val Asn Asn Pro
65 70 75 80

Cys Phe Gln Ser Pro Thr Pro Gly Glu Ser Pro Gly Ile Val Ser Asn
85 90 95

Tyr Asn Asn Ser Ile Leu Ala Arg Val Tyr Arg Asp Phe Gln Glu Leu
100 105 110

Leu Met Asn Ala Pro Glu Ser Gln His Leu Gly Arg Ile Trp Thr Glu
115 120 125

Leu His Ile Leu Ser Gln Phe Met Asp Thr Leu Arg Thr His Pro Glu
130 135 140

Arg Ile Ala Gly Arg Gly Ile Arg Ile Arg Asp Ile Leu Lys Asp Glu
145 150 155 160

Glu Thr Leu Thr Leu Phe Leu Ile Lys Asn Ile Gly Leu Ser Asp Ser
165 170 175

Val Val Tyr Leu Leu Ile Asn Ser Gln Val Arg Pro Glu Gln Phe Ala
180 185 190

His Gly Val Pro Asp Leu Ala Leu Lys Asp Ile Ala Cys Ser Glu Ala
195 200 205

```
Leu Leu Glu Arg Phe Ile Ile Phe Ser Gln Arg Arg Gly Ala Lys Thr
    210                 215                 220

Val Arg Tyr Ala Leu Cys Ser Leu Ser Gln Gly Thr Leu Gln Trp Ile
225                 230                 235                 240

Glu Asp Thr Leu Tyr Ala Asn Val Asp Phe Phe Lys Leu Phe Arg Val
                245                 250                 255

Leu Pro Thr Leu Leu Asp Ser Arg Ser Gln Gly Ile Asn Leu Arg Ser
            260                 265                 270

Trp Gly Gly Ile Leu Ser Asp Met Ser Pro Arg Ile Gln Glu Phe Ile
        275                 280                 285

His Arg Pro Ser Met Gln Asp Leu Leu Trp Val Thr Arg Pro Leu Met
    290                 295                 300

Gln Asn Gly Gly Pro Glu Thr Phe Thr Lys Leu Met Gly Ile Leu Ser
305                 310                 315                 320

Asp Leu Leu Cys Gly Tyr Pro Glu Gly Gly Gly Ser Arg Val Leu Ser
                325                 330                 335

Phe Asn Trp Tyr Glu Asp Asn Asn Tyr Lys Ala Phe Leu Gly Ile Asp
            340                 345                 350

Ser Thr Arg Lys Asp Pro Ile Tyr Ser Tyr Asp Arg Arg Thr Thr Ser
        355                 360                 365

Phe Cys Asn Ala Leu Ile Gln Ser Leu Glu Ser Asn Pro Leu Thr Lys
    370                 375                 380

Ile Ala Trp Arg Ala Ala Lys Pro Leu Leu Met Gly Lys Ile Leu Tyr
385                 390                 395                 400

Thr Pro Asp Ser Pro Ala Ala Arg Arg Ile Leu Lys Asn Ala Asn Ser
                405                 410                 415

Thr Phe Glu Glu Leu Glu His Val Arg Lys Leu Val Lys Ala Trp Glu
            420                 425                 430

Glu Val Gly Pro Gln Ile Trp Tyr Phe Phe Asp Asn Ser Thr Gln Met
        435                 440                 445
```

Asn Met Ile Arg Asp Thr Leu Gly Asn Pro Thr Val Lys Asp Phe Leu
450 455 460

Asn Arg Gln Leu Gly Glu Glu Gly Ile Thr Ala Glu Ala Ile Leu Asn
465 470 475 480

Phe Leu Tyr Lys Gly Pro Arg Glu Ser Gln Ala Asp Asp Met Ala Asn
485 490 495

Phe Asp Trp Arg Asp Ile Phe Asn Ile Thr Asp Arg Thr Leu Arg Leu
500 505 510

Val Asn Gln Tyr Leu Glu Cys Leu Val Leu Asp Lys Phe Glu Ser Tyr
515 520 525

Asn Asp Glu Thr Gln Leu Thr Gln Arg Ala Leu Ser Leu Leu Glu Glu
530 535 540

Asn Met Phe Trp Ala Gly Val Val Phe Pro Asp Met Tyr Pro Trp Thr
545 550 555 560

Ser Ser Leu Pro Pro His Val Lys Tyr Lys Ile Arg Met Asp Ile Asp
565 570 575

Val Val Glu Lys Thr Asn Lys Ile Lys Asp Arg Tyr Trp Asp Ser Gly
580 585 590

Pro Arg Ala Asp Pro Val Glu Asp Phe Arg Tyr Ile Trp Gly Gly Phe
595 600 605

Ala Tyr Leu Gln Asp Met Val Glu Gln Gly Ile Thr Arg Ser Gln Val
610 615 620

Gln Ala Glu Ala Pro Val Gly Ile Tyr Leu Gln Gln Met Pro Tyr Pro
625 630 635 640

Cys Phe Val Asp Asp Ser Phe Met Ile Ile Leu Asn Arg Cys Phe Pro
645 650 655

Ile Phe Met Val Leu Ala Trp Ile Tyr Ser Val Ser Met Thr Val Lys
660 665 670

Ser Ile Val Leu Glu Lys Glu Leu Arg Leu Lys Glu Thr Leu Lys Asn
675 680 685

Gln Gly Val Ser Asn Ala Val Ile Trp Cys Thr Trp Phe Leu Asp Ser

690 695 700

Phe Ser Ile Met Ser Met Ser Ile Phe Leu Leu Thr Ile Phe Ile Met
705 710 715 720

His Gly Arg Ile Leu His Tyr Ser Asp Pro Phe Ile Leu Phe Leu Phe
725 730 735

Leu Leu Ala Phe Ser Thr Ala Thr Ile Met Leu Cys Phe Leu Leu Ser
740 745 750

Thr Phe Phe Ser Lys Ala Ser Leu Ala Ala Ala Cys Ser Gly Val Ile
755 760 765

Tyr Phe Thr Leu Tyr Leu Pro His Ile Leu Cys Phe Ala Trp Gln Asp
770 775 780

Arg Met Thr Ala Glu Leu Lys Lys Ala Val Ser Leu Leu Ser Pro Val
785 790 795 800

Ala Phe Gly Phe Gly Thr Glu Tyr Leu Val Arg Phe Glu Glu Gln Gly
805 810 815

Leu Gly Leu Gln Trp Ser Asn Ile Gly Asn Ser Pro Thr Glu Gly Asp
820 825 830

Glu Phe Ser Phe Leu Leu Ser Met Gln Met Met Leu Leu Asp Ala Ala
835 840 845

Cys Tyr Gly Leu Leu Ala Trp Tyr Leu Asp Gln Val Phe Pro Gly Asp
850 855 860

Tyr Gly Thr Pro Leu Pro Trp Tyr Phe Leu Leu Gln Glu Ser Tyr Trp
865 870 875 880

Leu Ser Gly Glu Gly Cys Ser Thr Arg Glu Glu Arg Ala Leu Glu Lys
885 890 895

Thr Glu Pro Leu Thr Glu Glu Thr Glu Asp Pro Glu His Pro Glu Gly
900 905 910

Ile His Asp Ser Phe Phe Glu Arg Glu His Pro Gly Trp Val Pro Gly
915 920 925

Val Cys Val Lys Asn Leu Val Lys Ile Phe Glu Pro Cys Gly Arg Pro
930 935 940

```
Ala Val Asp Arg Leu Asn Ile Thr Phe Tyr Glu Asn Gln Ile Thr Ala
945                 950                 955                 960

Phe Leu Gly His Asn Gly Ala Gly Lys Thr Thr Thr Leu Ser Ile Leu
                965                 970                 975

Thr Gly Leu Leu Pro Pro Thr Ser Gly Thr Val Leu Val Gly Gly Arg
            980                 985                 990

Asp Ile Glu Thr Ser Leu Asp Ala  Val Arg Gln Ser Leu  Gly Met Cys
        995                 1000                1005

Pro Gln  His Asn Ile Leu Phe  His His Leu Thr Val  Ala Glu His
    1010                1015                1020

Met Leu  Phe Tyr Ala Gln Leu  Lys Gly Lys Ser Gln  Glu Glu Ala
    1025                1030                1035

Gln Leu  Glu Met Glu Ala Met  Leu Glu Asp Thr Gly  Leu His His
    1040                1045                1050

Lys Arg  Asn Glu Glu Ala Gln  Asp Leu Ser Gly Gly  Met Gln Arg
    1055                1060                1065

Lys Leu  Ser Val Ala Ile Ala  Phe Val Gly Asp Ala  Lys Val Val
    1070                1075                1080

Ile Leu  Asp Glu Pro Thr Ser  Gly Val Asp Pro Tyr  Ser Arg Arg
    1085                1090                1095

Ser Ile  Trp Asp Leu Leu Leu  Lys Tyr Arg Ser Gly  Arg Thr Ile
    1100                1105                1110

Ile Met  Pro Thr His His Met  Asp Glu Ala Asp His  Gln Gly Asp
    1115                1120                1125

Arg Ile  Ala Ile Ile Ala Gln  Gly Arg Leu Tyr Cys  Ser Gly Thr
    1130                1135                1140

Pro Leu  Phe Leu Lys Asn Cys  Phe Gly Thr Gly Leu  Tyr Leu Thr
    1145                1150                1155

Leu Val  Arg Lys Met Lys Asn  Ile Gln Ser Gln Arg  Lys Gly Ser
    1160                1165                1170
```

Glu Gly Thr Cys Ser Cys Ser Ser Lys Gly Phe Ser Thr Thr Cys
1175 1180 1185

Pro Ala His Val Asp Asp Leu Thr Pro Glu Gln Val Leu Asp Gly
1190 1195 1200

Asp Val Asn Glu Leu Met Asp Val Val Leu His His Val Pro Glu
1205 1210 1215

Ala Lys Leu Val Glu Cys Ile Gly Gln Glu Leu Ile Phe Leu Leu
1220 1225 1230

Pro Asn Lys Asn Phe Lys His Arg Ala Tyr Ala Ser Leu Phe Arg
1235 1240 1245

Glu Leu Glu Glu Thr Leu Ala Asp Leu Gly Leu Ser Ser Phe Gly
1250 1255 1260

Ile Ser Asp Thr Pro Leu Glu Glu Ile Phe Leu Lys Val Thr Glu
1265 1270 1275

Asp Ser Asp Ser Gly Pro Leu Phe Ala Gly Gly Ala Gln Gln Lys
1280 1285 1290

Arg Glu Asn Val Asn Pro Arg His Pro Cys Leu Gly Pro Arg Glu
1295 1300 1305

Lys Ala Gly Gln Thr Pro Gln Asp Ser Asn Val Cys Ser Pro Gly
1310 1315 1320

Ala Pro Ala Ala His Pro Glu Gly Gln Pro Pro Pro Glu Pro Glu
1325 1330 1335

Cys Pro Gly Pro Gln Leu Asn Thr Gly Thr Gln Leu Val Leu Gln
1340 1345 1350

His Val Gln Ala Leu Leu Val Lys Arg Phe Gln His Thr Ile Arg
1355 1360 1365

Ser His Lys Asp Phe Leu Ala Gln Ile Val Leu Pro Ala Thr Phe
1370 1375 1380

Val Phe Leu Ala Leu Met Leu Ser Ile Val Ile Leu Pro Phe Gly
1385 1390 1395

Glu Tyr Pro Ala Leu Thr Leu His Pro Trp Ile Tyr Gly Gln Gln
1400 1405 1410

Tyr Thr Phe Phe Ser Met Asp Glu Pro Gly Ser Glu Gln Phe Thr
1415 1420 1425

Val Leu Ala Asp Val Leu Leu Asn Lys Pro Gly Phe Gly Asn Arg
1430 1435 1440

Cys Leu Lys Glu Gly Trp Leu Pro Glu Tyr Pro Cys Gly Asn Ser
1445 1450 1455

Thr Pro Trp Lys Thr Pro Ser Val Ser Pro Asn Ile Thr Gln Leu
1460 1465 1470

Phe Gln Lys Gln Lys Trp Thr Gln Val Asn Pro Ser Pro Ser Cys
1475 1480 1485

Arg Cys Ser Thr Arg Glu Lys Leu Thr Met Leu Pro Glu Cys Pro
1490 1495 1500

Glu Gly Ala Gly Gly Leu Pro Pro Pro Gln Arg Thr Gln Arg Ser
1505 1510 1515

Thr Glu Ile Leu Gln Asp Leu Thr Asp Arg Asn Ile Ser Asp Phe
1520 1525 1530

Leu Val Lys Thr Tyr Pro Ala Leu Ile Arg Ser Ser Leu Lys Ser
1535 1540 1545

Lys Phe Trp Val Asn Glu Gln Arg Tyr Gly Gly Ile Ser Ile Gly
1550 1555 1560

Gly Lys Leu Pro Val Val Pro Ile Thr Gly Glu Ala Leu Val Gly
1565 1570 1575

Phe Leu Ser Asp Leu Gly Arg Ile Met Asn Val Ser Gly Gly Pro
1580 1585 1590

Ile Thr Arg Glu Ala Ser Lys Glu Ile Pro Asp Phe Leu Lys His
1595 1600 1605

Leu Glu Thr Glu Asp Asn Ile Lys Val Trp Phe Asn Asn Lys Gly
1610 1615 1620

Trp His Ala Leu Val Ser Phe Leu Asn Val Ala His Asn Ala Ile

1625 1630 1635

Leu Arg Ala Ser Leu Pro Lys Asp Arg Ser Pro Glu Glu Tyr Gly
1640 1645 1650

Ile Thr Val Ile Ser Gln Pro Leu Asn Leu Thr Lys Glu Gln Leu
1655 1660 1665

Ser Glu Ile Thr Val Leu Thr Thr Ser Val Asp Ala Val Val Ala
1670 1675 1680

Ile Cys Val Ile Phe Ser Met Ser Phe Val Pro Ala Ser Phe Val
1685 1690 1695

Leu Tyr Leu Ile Gln Glu Arg Val Asn Lys Ser Lys His Leu Gln
1700 1705 1710

Phe Ile Ser Gly Val Ser Pro Thr Thr Tyr Trp Val Thr Asn Phe
1715 1720 1725

Leu Trp Asp Ile Met Asn Tyr Ser Val Ser Ala Gly Leu Val Val
1730 1735 1740

Gly Ile Phe Ile Gly Phe Gln Lys Lys Ala Tyr Thr Ser Pro Glu
1745 1750 1755

Asn Leu Pro Ala Leu Val Ala Leu Leu Leu Leu Tyr Gly Trp Ala
1760 1765 1770

Val Ile Pro Met Met Tyr Pro Ala Ser Phe Leu Phe Asp Val Pro
1775 1780 1785

Ser Thr Ala Tyr Val Ala Leu Ser Cys Ala Asn Leu Phe Ile Gly
1790 1795 1800

Ile Asn Ser Ser Ala Ile Thr Phe Ile Leu Glu Leu Phe Asp Asn
1805 1810 1815

Asn Arg Thr Leu Leu Arg Phe Asn Ala Val Leu Arg Lys Leu Leu
1820 1825 1830

Ile Val Phe Pro His Phe Cys Leu Gly Arg Gly Leu Ile Asp Leu
1835 1840 1845

Ala Leu Ser Gln Ala Val Thr Asp Val Tyr Ala Arg Phe Gly Glu
1850 1855 1860

Glu His Ser Ala Asn Pro Phe His Trp Asp Leu Ile Gly Lys Asn
1865 1870 1875

Leu Phe Ala Met Val Val Glu Gly Val Val Tyr Phe Leu Leu Thr
1880 1885 1890

Leu Leu Val Gln Arg His Phe Phe Leu Ser Gln Trp Ile Ala Glu
1895 1900 1905

Pro Thr Lys Glu Pro Ile Val Asp Glu Asp Asp Asp Val Ala Glu
1910 1915 1920

Glu Arg Gln Arg Ile Ile Thr Gly Gly Asn Lys Thr Asp Ile Leu
1925 1930 1935

Arg Leu His Glu Leu Thr Lys Ile Tyr Leu Gly Thr Ser Ser Pro
1940 1945 1950

Ala Val Asp Arg Leu Cys Val Gly Val Arg Pro Gly Glu Cys Phe
1955 1960 1965

Gly Leu Leu Gly Val Asn Gly Ala Gly Lys Thr Thr Thr Phe Lys
1970 1975 1980

Met Leu Thr Gly Asp Thr Thr Val Thr Ser Gly Asp Ala Thr Val
1985 1990 1995

Ala Gly Lys Ser Ile Leu Thr Asn Ile Ser Glu Val His Gln Asn
2000 2005 2010

Met Gly Tyr Cys Pro Gln Phe Asp Ala Ile Asp Glu Leu Leu Thr
2015 2020 2025

Gly Arg Glu His Leu Tyr Leu Tyr Ala Arg Leu Arg Gly Val Pro
2030 2035 2040

Ala Glu Glu Ile Glu Lys Val Ala Asn Trp Ser Ile Lys Ser Leu
2045 2050 2055

Gly Leu Thr Val Tyr Ala Asp Cys Leu Ala Gly Thr Tyr Ser Gly
2060 2065 2070

Gly Asn Lys Arg Lys Leu Ser Thr Ala Ile Ala Leu Ile Gly Cys
2075 2080 2085

Pro Pro Leu Val Leu Leu Asp Glu Pro Thr Thr Gly Met Asp Pro
2090 2095 2100

Gln Ala Arg Arg Met Leu Trp Asn Val Ile Val Ser Ile Ile Arg
2105 2110 2115

Lys Gly Arg Ala Val Val Leu Thr Ser His Ser Met Glu Glu Cys
2120 2125 2130

Glu Ala Leu Cys Thr Arg Leu Ala Ile Met Val Lys Gly Ala Phe
2135 2140 2145

Arg Cys Met Gly Thr Ile Gln His Leu Lys Ser Lys Phe Gly Asp
2150 2155 2160

Gly Tyr Ile Val Thr Met Lys Ile Lys Ser Pro Lys Asp Asp Leu
2165 2170 2175

Leu Pro Asp Leu Asn Pro Val Glu Gln Phe Phe Gln Gly Asn Phe
2180 2185 2190

Pro Gly Ser Val Gln Arg Glu Arg His Tyr Asn Met Leu Gln Phe
2195 2200 2205

Gln Val Ser Ser Ser Ser Leu Ala Arg Ile Phe Gln Leu Leu Leu
2210 2215 2220

Ser His Lys Asp Ser Leu Leu Ile Glu Glu Tyr Ser Val Thr Gln
2225 2230 2235

Thr Thr Leu Asp Gln Val Phe Val Asn Phe Ala Lys Gln Gln Thr
2240 2245 2250

Glu Ser His Asp Leu Pro Leu His Pro Arg Ala Ala Gly Ala Ser
2255 2260 2265

Arg Gln Ala Gln Asp
2270

<210> 71
<211> 560
<212> PRT
<213> Homo Sapiens

<400> 71

Met Val Pro His Ala Ile Leu Ala Arg Gly Arg Asp Val Cys Arg Arg

1 5 10 15

Asn Gly Leu Leu Ile Leu Ser Val Leu Ser Val Ile Val Gly Cys Leu
20 25 30

Leu Gly Phe Phe Leu Arg Thr Arg Arg Leu Ser Pro Gln Glu Ile Ser
35 40 45

Tyr Phe Gln Phe Pro Gly Glu Leu Leu Met Arg Met Leu Lys Met Met
50 55 60

Ile Leu Pro Leu Val Val Ser Ser Leu Met Ser Gly Leu Ala Ser Leu
65 70 75 80

Asp Ala Lys Thr Ser Ser Arg Leu Gly Val Leu Thr Val Ala Tyr Tyr
85 90 95

Leu Trp Thr Thr Phe Met Ala Val Ile Val Gly Ile Phe Met Val Ser
100 105 110

Ile Ile His Pro Gly Ser Ala Ala Gln Lys Glu Thr Thr Glu Gln Ser
115 120 125

Gly Lys Pro Ile Met Ser Ser Ala Asp Ala Leu Leu Asp Leu Ile Arg
130 135 140

Asn Met Phe Pro Ala Asn Leu Val Glu Ala Thr Phe Lys Gln Tyr Arg
145 150 155 160

Thr Lys Thr Thr Pro Val Val Lys Ser Pro Lys Val Ala Pro Glu Glu
165 170 175

Ala Pro Pro Arg Arg Ile Leu Ile Tyr Gly Val Gln Glu Glu Asn Gly
180 185 190

Ser His Val Gln Asn Phe Ala Leu Asp Leu Thr Pro Pro Pro Glu Val
195 200 205

Val Tyr Lys Ser Glu Pro Gly Thr Ser Asp Gly Met Asn Val Leu Gly
210 215 220

Ile Val Phe Phe Ser Ala Thr Met Gly Ile Met Leu Gly Arg Met Gly
225 230 235 240

Asp Ser Gly Ala Pro Leu Val Ser Phe Cys Gln Cys Leu Asn Glu Ser
245 250 255

Val Met Lys Ile Val Ala Val Ala Val Trp Tyr Phe Pro Phe Gly Ile
260 265 270

Val Phe Leu Ile Ala Gly Lys Ile Leu Glu Met Asp Asp Pro Arg Ala
275 280 285

Val Gly Lys Lys Leu Gly Phe Tyr Ser Val Thr Val Val Cys Gly Leu
290 295 300

Val Leu His Gly Leu Phe Ile Leu Pro Leu Leu Tyr Phe Phe Ile Thr
305 310 315 320

Lys Lys Asn Pro Ile Val Phe Ile Arg Gly Ile Leu Gln Ala Leu Leu
325 330 335

Ile Ala Leu Ala Thr Ser Ser Ser Ser Ala Thr Leu Pro Ile Thr Phe
340 345 350

Lys Cys Leu Leu Glu Asn Asn His Ile Asp Arg Arg Ile Ala Arg Phe
355 360 365

Val Leu Pro Val Gly Ala Thr Ile Asn Met Asp Gly Thr Ala Leu Tyr
370 375 380

Glu Ala Val Ala Ala Ile Phe Ile Ala Gln Val Asn Asn Tyr Glu Leu
385 390 395 400

Asp Phe Gly Gln Ile Ile Thr Ile Ser Ile Thr Ala Thr Ala Ala Ser
405 410 415

Ile Gly Ala Ala Gly Ile Pro Gln Ala Gly Leu Val Thr Met Val Ile
420 425 430

Val Leu Thr Ser Val Gly Leu Pro Thr Asp Asp Ile Thr Leu Ile Ile
435 440 445

Ala Val Asp Trp Ala Leu Asp Arg Phe Arg Thr Met Ile Asn Val Leu
450 455 460

Gly Asp Ala Leu Ala Ala Gly Ile Met Ala His Ile Cys Arg Lys Asp
465 470 475 480

Phe Ala Arg Asp Thr Gly Thr Glu Lys Leu Leu Pro Cys Glu Thr Lys
485 490 495

Pro Val Ser Leu Gln Glu Ile Val Ala Ala Gln Gln Asn Gly Cys Val
500 505 510

Lys Ser Val Ala Glu Ala Ser Glu Leu Thr Leu Gly Pro Thr Cys Pro
515 520 525

His His Val Pro Val Gln Val Glu Arg Asp Glu Glu Leu Pro Ala Ala
530 535 540

Ser Leu Asn His Cys Thr Ile Gln Ile Ser Glu Leu Glu Thr Asn Val
545 550 555 560

<210> 72
<211> 840
<212> PRT
<213> Homo Sapiens

<400> 72

Met Val Thr Val Gly Asn Tyr Cys Glu Ala Glu Gly Pro Val Gly Pro
1 5 10 15

Ala Trp Met Gln Asp Gly Leu Ser Pro Cys Phe Phe Phe Thr Leu Val
20 25 30

Pro Ser Thr Arg Met Ala Leu Gly Thr Leu Ala Leu Val Leu Ala Leu
35 40 45

Pro Cys Arg Arg Arg Glu Arg Pro Ala Gly Ala Asp Ser Leu Ser Trp
50 55 60

Gly Ala Gly Pro Arg Ile Ser Pro Tyr Val Leu Gln Leu Leu Leu Ala
65 70 75 80

Thr Leu Gln Ala Ala Leu Pro Leu Ala Gly Leu Ala Gly Arg Val Gly
85 90 95

Thr Ala Arg Gly Ala Pro Leu Pro Ser Tyr Leu Leu Leu Ala Ser Val
100 105 110

Leu Glu Ser Leu Ala Gly Ala Cys Gly Leu Trp Leu Leu Val Val Glu
115 120 125

Arg Ser Gln Ala Arg Gln Arg Leu Ala Met Gly Ile Trp Ile Lys Phe
130 135 140

Arg His Ser Pro Gly Leu Leu Leu Leu Trp Thr Val Ala Phe Ala Ala

```
145                 150                 155                 160

Glu Asn Leu Ala Leu Val Ser Trp Asn Ser Pro Gln Trp Trp Trp Ala
                165                 170                 175

Arg Ala Asp Leu Gly Gln Gln Val Gln Phe Ser Leu Trp Val Leu Arg
            180                 185                 190

Tyr Val Val Ser Gly Gly Leu Phe Val Leu Gly Leu Trp Ala Pro Gly
        195                 200                 205

Leu Arg Pro Gln Ser Tyr Thr Leu Gln Val His Glu Glu Asp Gln Asp
    210                 215                 220

Val Glu Arg Ser Gln Val Arg Ser Ala Ala Gln Gln Ser Thr Trp Arg
225                 230                 235                 240

Asp Phe Gly Arg Lys Leu Arg Leu Leu Ser Gly Tyr Leu Trp Pro Arg
                245                 250                 255

Gly Ser Pro Ala Leu Gln Leu Val Val Leu Ile Cys Leu Gly Leu Met
            260                 265                 270

Gly Leu Glu Arg Ala Leu Asn Val Leu Val Pro Ile Phe Tyr Arg Asn
        275                 280                 285

Ile Val Asn Leu Leu Thr Glu Lys Ala Pro Trp Asn Ser Leu Ala Trp
    290                 295                 300

Thr Val Thr Ser Tyr Val Phe Leu Lys Phe Leu Gln Gly Gly Gly Thr
305                 310                 315                 320

Gly Ser Thr Gly Phe Val Ser Asn Leu Arg Thr Phe Leu Trp Ile Arg
                325                 330                 335

Val Gln Gln Phe Thr Ser Arg Arg Val Glu Leu Leu Ile Phe Ser His
            340                 345                 350

Leu His Glu Leu Ser Leu Arg Trp His Leu Gly Arg Arg Thr Gly Glu
        355                 360                 365

Val Leu Arg Ile Ala Asp Arg Gly Thr Ser Ser Val Thr Gly Leu Leu
    370                 375                 380

Ser Tyr Leu Val Phe Asn Val Ile Pro Thr Leu Ala Asp Ile Ile Ile
385                 390                 395                 400
```

Gly Ile Ile Tyr Phe Ser Met Phe Phe Asn Ala Trp Phe Gly Leu Ile
405 410 415

Val Phe Leu Cys Met Ser Leu Tyr Leu Thr Leu Thr Ile Val Val Thr
420 425 430

Glu Trp Arg Thr Lys Phe Arg Arg Ala Met Asn Thr Gln Glu Asn Ala
435 440 445

Thr Arg Ala Arg Ala Val Asp Ser Leu Leu Asn Phe Glu Thr Val Lys
450 455 460

Tyr Tyr Asn Ala Glu Ser Tyr Glu Val Glu Arg Tyr Arg Glu Ala Ile
465 470 475 480

Ile Lys Tyr Gln Gly Leu Glu Trp Lys Ser Ser Ala Ser Leu Val Leu
485 490 495

Leu Asn Gln Thr Gln Asn Leu Val Ile Gly Leu Gly Leu Leu Ala Gly
500 505 510

Ser Leu Leu Cys Ala Tyr Phe Val Thr Glu Gln Lys Leu Gln Val Gly
515 520 525

Asp Tyr Val Leu Phe Gly Thr Tyr Ile Ile Gln Leu Tyr Met Pro Leu
530 535 540

Asn Trp Phe Gly Thr Tyr Tyr Arg Met Ile Gln Thr Asn Phe Ile Asp
545 550 555 560

Met Glu Asn Met Phe Asp Leu Leu Lys Glu Glu Thr Glu Val Lys Asp
565 570 575

Leu Pro Gly Ala Gly Pro Leu Arg Phe Gln Lys Gly Arg Ile Glu Phe
580 585 590

Glu Asn Val His Phe Ser Tyr Ala Asp Gly Arg Glu Thr Leu Gln Asp
595 600 605

Val Ser Phe Thr Val Met Pro Gly Gln Thr Leu Ala Leu Val Gly Pro
610 615 620

Ser Gly Ala Gly Lys Ser Thr Ile Leu Arg Leu Leu Phe Arg Phe Tyr
625 630 635 640

Asp Ile Ser Ser Gly Cys Ile Arg Ile Asp Gly Gln Asp Ile Ser Gln
645 650 655

Val Thr Gln Ala Ser Leu Arg Ser His Ile Gly Val Val Pro Gln Asp
660 665 670

Thr Val Leu Phe Asn Asp Thr Ile Ala Asp Asn Ile Arg Tyr Gly Arg
675 680 685

Val Thr Ala Gly Asn Asp Glu Val Glu Ala Ala Ala Gln Ala Ala Gly
690 695 700

Ile His Asp Ala Ile Met Ala Phe Pro Glu Gly Tyr Arg Thr Gln Val
705 710 715 720

Gly Glu Arg Gly Leu Lys Leu Ser Gly Gly Glu Lys Gln Arg Val Ala
725 730 735

Ile Ala Arg Thr Ile Leu Lys Ala Pro Gly Ile Ile Leu Leu Asp Glu
740 745 750

Ala Thr Ser Ala Leu Asp Thr Ser Asn Glu Arg Ala Ile Gln Ala Ser
755 760 765

Leu Ala Lys Val Cys Ala Asn Arg Thr Thr Ile Val Val Ala His Arg
770 775 780

Leu Ser Thr Val Val Asn Ala Asp Gln Ile Leu Val Ile Lys Asp Gly
785 790 795 800

Cys Ile Val Glu Arg Gly Arg His Glu Ala Leu Leu Ser Arg Gly Gly
805 810 815

Val Tyr Ala Asp Met Trp Gln Leu Gln Gln Gly Gln Glu Glu Thr Ser
820 825 830

Glu Asp Thr Lys Pro Gln Thr Met
835 840

<210> 73
<211> 332
<212> PRT
<213> Homo Sapiens

<400> 73

Met Leu Leu Glu Thr Gln Asp Ala Leu Tyr Val Ala Leu Glu Leu Val

1 5 10 15

Ile Ala Ala Leu Ser Val Ala Gly Asn Val Leu Val Cys Ala Ala Val
20 25 30

Gly Thr Ala Asn Thr Leu Gln Thr Pro Thr Asn Tyr Phe Leu Val Ser
35 40 45

Leu Ala Ala Ala Asp Val Ala Val Gly Leu Phe Ala Ile Pro Phe Ala
50 55 60

Ile Thr Ile Ser Leu Gly Phe Cys Thr Asp Phe Tyr Gly Cys Leu Phe
65 70 75 80

Leu Ala Cys Phe Val Leu Val Leu Thr Gln Ser Ser Ile Phe Ser Leu
85 90 95

Leu Ala Val Ala Val Asp Arg Tyr Leu Ala Ile Cys Val Pro Leu Arg
100 105 110

Tyr Lys Ser Leu Val Thr Gly Thr Arg Ala Arg Gly Val Ile Ala Val
115 120 125

Leu Trp Val Leu Ala Phe Gly Ile Gly Leu Thr Pro Phe Leu Gly Trp
130 135 140

Asn Ser Lys Asp Ser Ala Thr Asn Asn Cys Thr Glu Pro Trp Asp Gly
145 150 155 160

Thr Thr Asn Glu Ser Cys Cys Leu Val Lys Cys Leu Phe Glu Asn Val
165 170 175

Val Pro Met Ser Tyr Met Val Tyr Phe Asn Phe Phe Gly Cys Val Leu
180 185 190

Pro Pro Leu Leu Ile Met Leu Val Ile Tyr Ile Lys Ile Phe Leu Val
195 200 205

Ala Cys Arg Gln Leu Gln Arg Thr Glu Leu Met Asp His Ser Arg Thr
210 215 220

Thr Leu Gln Arg Glu Ile His Ala Ala Lys Ser Leu Ala Met Ile Val
225 230 235 240

Gly Ile Phe Ala Leu Cys Trp Leu Pro Val His Ala Val Asn Cys Val
245 250 255

Thr Leu Phe Gln Pro Ala Gln Gly Lys Asn Lys Pro Lys Trp Ala Met
260 265 270

Asn Met Ala Ile Leu Leu Ser His Ala Asn Ser Val Val Asn Pro Ile
275 280 285

Val Tyr Ala Tyr Arg Asn Arg Asp Phe Arg Tyr Thr Phe His Lys Ile
290 295 300

Ile Ser Arg Tyr Leu Leu Cys Gln Ala Asp Val Lys Ser Gly Asn Gly
305 310 315 320

Gln Ala Gly Val Gln Pro Ala Leu Gly Val Gly Leu
325 330

<210> 74
<211> 180
<212> PRT
<213> Homo Sapiens

<400> 74

Met Gly Leu Gly Ala Arg Gly Ala Trp Ala Ala Leu Leu Leu Gly Thr
1 5 10 15

Leu Gln Val Leu Ala Leu Leu Gly Ala Ala His Glu Ser Ala Ala Met
20 25 30

Ala Glu Thr Leu Gln His Val Pro Ser Asp His Thr Asn Glu Thr Ser
35 40 45

Asn Ser Thr Val Lys Pro Pro Thr Ser Val Ala Ser Asp Ser Ser Asn
50 55 60

Thr Thr Val Thr Thr Met Lys Pro Thr Ala Ala Ser Asn Thr Thr Thr
65 70 75 80

Pro Gly Met Val Ser Thr Asn Met Thr Ser Thr Thr Leu Lys Ser Thr
85 90 95

Pro Lys Thr Thr Ser Val Ser Gln Asn Thr Ser Gln Ile Ser Thr Ser
100 105 110

Thr Met Thr Val Thr His Asn Ser Ser Val Thr Ser Ala Ala Ser Ser
115 120 125

Val Thr Ile Thr Thr Thr Met His Ser Glu Ala Lys Lys Gly Ser Lys
130 135 140

Phe Asp Thr Gly Ser Phe Val Gly Gly Ile Val Leu Thr Leu Gly Val
145 150 155 160

Leu Ser Ile Leu Tyr Ile Gly Cys Lys Met Tyr Tyr Ser Arg Arg Gly
165 170 175

Ile Arg Tyr Arg
180

<210> 75
<211> 240
<212> PRT
<213> Homo Sapiens

<400> 75

Met Ala Gln His Gly Ala Met Gly Ala Phe Arg Ala Leu Cys Gly Leu
1 5 10 15

Ala Leu Leu Cys Ala Leu Ser Leu Gly Gln Arg Pro Thr Gly Gly Pro
20 25 30

Gly Cys Gly Pro Gly Arg Leu Leu Leu Gly Thr Gly Thr Asp Ala Arg
35 40 45

Cys Cys Arg Val His Thr Thr Arg Cys Cys Arg Asp Tyr Pro Gly Glu
50 55 60

Glu Cys Cys Ser Glu Trp Asp Cys Met Cys Val Gln Pro Glu Phe His
65 70 75 80

Cys Gly Asp Pro Cys Cys Thr Thr Cys Arg His His Pro Cys Pro Pro
85 90 95

Gly Gln Gly Val Gln Ser Gln Gly Lys Phe Ser Phe Gly Phe Gln Cys
100 105 110

Ile Asp Cys Ala Ser Gly Thr Phe Ser Gly Gly His Glu Gly His Cys
115 120 125

Lys Pro Trp Thr Asp Cys Thr Gln Phe Gly Phe Leu Thr Val Phe Pro
130 135 140

Gly Asn Lys Thr His Asn Ala Val Cys Val Pro Gly Ser Pro Pro Ala
145 150 155 160

```
Glu Pro Leu Gly Trp Leu Thr Val Val Leu Leu Ala Val Ala Ala Cys
                165                 170                 175

Val Leu Leu Leu Thr Ser Ala Gln Leu Gly Leu His Ile Trp Gln Leu
            180                 185                 190

Arg Ser Gln Cys Met Trp Pro Arg Glu Thr Gln Leu Leu Leu Glu Val
        195                 200                 205

Pro Pro Ser Thr Glu Asp Ala Arg Ser Cys Gln Phe Pro Glu Glu Glu
    210                 215                 220

Arg Gly Glu Arg Ser Ala Glu Glu Lys Gly Arg Leu Gly Asp Leu Trp
225                 230                 235                 240
```

<210> 76
<211> 514
<212> PRT
<213> Homo Sapiens

<400> 76

```
Met Gly Cys Asp Gly Arg Val Ser Gly Leu Leu Arg Arg Asn Leu Gln
1               5                   10                  15

Pro Thr Leu Thr Tyr Trp Ser Val Phe Phe Ser Phe Gly Leu Cys Ile
            20                  25                  30

Ala Phe Leu Gly Pro Thr Leu Leu Asp Leu Arg Cys Gln Thr His Ser
        35                  40                  45

Ser Leu Pro Gln Ile Ser Trp Val Phe Phe Ser Gln Gln Leu Cys Leu
    50                  55                  60

Leu Leu Gly Ser Ala Leu Gly Gly Val Phe Lys Arg Thr Leu Ala Gln
65                  70                  75                  80

Ser Leu Trp Ala Leu Phe Thr Ser Ser Leu Ala Ile Ser Leu Val Phe
                85                  90                  95

Ala Val Ile Pro Phe Cys Arg Asp Val Lys Val Leu Ala Ser Val Met
            100                 105                 110

Ala Leu Ala Gly Leu Ala Met Gly Cys Ile Asp Thr Val Ala Asn Met
        115                 120                 125
```

Gln Leu Val Arg Met Tyr Gln Lys Asp Ser Ala Val Phe Leu Gln Val
130 135 140

Leu His Phe Phe Val Gly Phe Gly Ala Leu Leu Ser Pro Leu Ile Ala
145 150 155 160

Asp Pro Phe Leu Ser Glu Ala Asn Cys Leu Pro Ala Asn Ser Thr Ala
165 170 175

Asn Thr Thr Ser Arg Gly His Leu Phe His Val Ser Arg Val Leu Gly
180 185 190

Gln His His Val Asp Ala Lys Pro Trp Ser Asn Gln Thr Phe Pro Gly
195 200 205

Leu Thr Pro Lys Asp Gly Ala Gly Thr Arg Val Ser Tyr Ala Phe Trp
210 215 220

Ile Met Ala Leu Ile Asp Leu Pro Val Pro Met Ala Val Leu Met Leu
225 230 235 240

Leu Ser Lys Glu Arg Leu Leu Thr Cys Cys Pro Gln Arg Arg Pro Leu
245 250 255

Leu Leu Ser Ala Asp Glu Leu Ala Leu Glu Thr Gln Pro Pro Glu Lys
260 265 270

Glu Asp Ala Ser Ser Leu Pro Pro Lys Phe Gln Ser His Leu Gly His
275 280 285

Glu Asp Leu Phe Ser Cys Cys Gln Arg Lys Asn Leu Arg Gly Ala Pro
290 295 300

Tyr Ser Phe Phe Ala Ile His Ile Thr Gly Ala Leu Val Leu Phe Met
305 310 315 320

Thr Asp Gly Leu Thr Gly Ala Tyr Ser Ala Phe Val Tyr Ser Tyr Ala
325 330 335

Val Glu Lys Pro Leu Ser Val Gly His Lys Val Ala Gly Tyr Leu Pro
340 345 350

Ser Leu Phe Trp Gly Phe Ile Thr Leu Gly Arg Leu Leu Ser Ile Pro
355 360 365

Ile Ser Ser Arg Met Lys Pro Ala Thr Met Val Phe Ile Asn Val Val

370 375 380

Gly Val Val Val Thr Phe Leu Val Leu Leu Ile Phe Ser Tyr Asn Val
385 390 395 400

Val Phe Leu Phe Val Gly Thr Ala Ser Leu Gly Leu Phe Leu Ser Ser
405 410 415

Thr Phe Pro Ser Met Leu Ala Tyr Thr Glu Asp Ser Leu Gln Tyr Lys
420 425 430

Gly Cys Ala Thr Thr Val Leu Val Thr Gly Ala Gly Val Gly Glu Met
435 440 445

Val Leu Gln Met Leu Val Gly Ser Ile Phe Gln Ala Gln Gly Ser Tyr
450 455 460

Ser Phe Leu Val Cys Gly Val Ile Phe Gly Cys Leu Ala Phe Thr Phe
465 470 475 480

Tyr Ile Leu Leu Leu Phe Phe His Arg Met His Pro Gly Leu Pro Ser
485 490 495

Val Pro Thr Gln Asp Arg Ser Ile Gly Met Glu Asn Ser Glu Cys Tyr
500 505 510

Gln Arg

<210> 77
<211> 1181
<212> PRT
<213> Homo Sapiens

<400> 77

Met Gly Pro Glu Arg Thr Gly Ala Ala Pro Leu Pro Leu Leu Leu Val
1 5 10 15

Leu Ala Leu Ser Gln Gly Ile Leu Asn Cys Cys Leu Ala Tyr Asn Val
20 25 30

Gly Leu Pro Glu Ala Lys Ile Phe Ser Gly Pro Ser Ser Glu Gln Phe
35 40 45

Gly Tyr Ala Val Gln Gln Phe Ile Asn Pro Lys Gly Asn Trp Leu Leu
50 55 60

Val Gly Ser Pro Trp Ser Gly Phe Pro Glu Asn Arg Met Gly Asp Val
65 70 75 80

Tyr Lys Cys Pro Val Asp Leu Ser Thr Ala Thr Cys Glu Lys Leu Asn
85 90 95

Leu Gln Thr Ser Thr Ser Ile Pro Asn Val Thr Glu Met Lys Thr Asn
100 105 110

Met Ser Leu Gly Leu Ile Leu Thr Arg Asn Met Gly Thr Gly Gly Phe
115 120 125

Leu Thr Cys Gly Pro Leu Trp Ala Gln Gln Cys Gly Asn Gln Tyr Tyr
130 135 140

Thr Thr Gly Val Cys Ser Asp Ile Ser Pro Asp Phe Gln Leu Ser Ala
145 150 155 160

Ser Phe Ser Pro Ala Thr Gln Pro Cys Pro Ser Leu Ile Asp Val Val
165 170 175

Val Val Cys Asp Glu Ser Asn Ser Ile Tyr Pro Trp Asp Ala Val Lys
180 185 190

Asn Phe Leu Glu Lys Phe Val Gln Gly Leu Asp Ile Gly Pro Thr Lys
195 200 205

Thr Gln Val Gly Leu Ile Gln Tyr Ala Asn Asn Pro Arg Val Val Phe
210 215 220

Asn Leu Asn Thr Tyr Lys Thr Lys Glu Glu Met Ile Val Ala Thr Ser
225 230 235 240

Gln Thr Ser Gln Tyr Gly Gly Asp Leu Thr Asn Thr Phe Gly Ala Ile
245 250 255

Gln Tyr Ala Arg Lys Tyr Ala Tyr Ser Ala Ala Ser Gly Gly Arg Arg
260 265 270

Ser Ala Thr Lys Val Met Val Val Val Thr Asp Gly Glu Ser His Asp
275 280 285

Gly Ser Met Leu Lys Ala Val Ile Asp Gln Cys Asn His Asp Asn Ile
290 295 300

Leu Arg Phe Gly Ile Ala Val Leu Gly Tyr Leu Asn Arg Asn Ala Leu
305 310 315 320

Asp Thr Lys Asn Leu Ile Lys Glu Ile Lys Ala Ile Ala Ser Ile Pro
325 330 335

Thr Glu Arg Tyr Phe Phe Asn Val Ser Asp Glu Ala Ala Leu Leu Glu
340 345 350

Lys Ala Gly Thr Leu Gly Glu Gln Ile Phe Ser Ile Glu Gly Thr Val
355 360 365

Gln Gly Gly Asp Asn Phe Gln Met Glu Met Ser Gln Val Gly Phe Ser
370 375 380

Ala Asp Tyr Ser Ser Gln Asn Asp Ile Leu Met Leu Gly Ala Val Gly
385 390 395 400

Ala Phe Gly Trp Ser Gly Thr Ile Val Gln Lys Thr Ser His Gly His
405 410 415

Leu Ile Phe Pro Lys Gln Ala Phe Asp Gln Ile Leu Gln Asp Arg Asn
420 425 430

His Ser Ser Tyr Leu Gly Tyr Ser Val Ala Ala Ile Ser Thr Gly Glu
435 440 445

Ser Thr His Phe Val Ala Gly Ala Pro Arg Ala Asn Tyr Thr Gly Gln
450 455 460

Ile Val Leu Tyr Ser Val Asn Glu Asn Gly Asn Ile Thr Val Ile Gln
465 470 475 480

Ala His Arg Gly Asp Gln Ile Gly Ser Tyr Phe Gly Ser Val Leu Cys
485 490 495

Ser Val Asp Val Asp Lys Asp Thr Ile Thr Asp Val Leu Leu Val Gly
500 505 510

Ala Pro Met Tyr Met Ser Asp Leu Lys Lys Glu Glu Gly Arg Val Tyr
515 520 525

Leu Phe Thr Ile Lys Lys Gly Ile Leu Gly Gln His Gln Phe Leu Glu
530 535 540

Gly Pro Glu Gly Ile Glu Asn Thr Arg Phe Gly Ser Ala Ile Ala Ala

545 550 555 560

Leu Ser Asp Ile Asn Met Asp Gly Phe Asn Asp Val Ile Val Gly Ser
565 570 575

Pro Leu Glu Asn Gln Asn Ser Gly Ala Val Tyr Ile Tyr Asn Gly His
580 585 590

Gln Gly Thr Ile Arg Thr Lys Tyr Ser Gln Lys Ile Leu Gly Ser Asp
595 600 605

Gly Ala Phe Arg Ser His Leu Gln Tyr Phe Gly Arg Ser Leu Asp Gly
610 615 620

Tyr Gly Asp Leu Asn Gly Asp Ser Ile Thr Asp Val Ser Ile Gly Ala
625 630 635 640

Phe Gly Gln Val Val Gln Leu Trp Ser Gln Ser Ile Ala Asp Val Ala
645 650 655

Ile Glu Ala Ser Phe Thr Pro Glu Lys Ile Thr Leu Val Asn Lys Asn
660 665 670

Ala Gln Ile Ile Leu Lys Leu Cys Phe Ser Ala Lys Phe Arg Pro Thr
675 680 685

Lys Gln Asn Asn Gln Val Ala Ile Val Tyr Asn Ile Thr Leu Asp Ala
690 695 700

Asp Gly Phe Ser Ser Arg Val Thr Ser Arg Gly Leu Phe Lys Glu Asn
705 710 715 720

Asn Glu Arg Cys Leu Gln Lys Asn Met Val Val Asn Gln Ala Gln Ser
725 730 735

Cys Pro Glu His Ile Ile Tyr Ile Gln Glu Pro Ser Asp Val Val Asn
740 745 750

Ser Leu Asp Leu Arg Val Asp Ile Ser Leu Glu Asn Pro Gly Thr Ser
755 760 765

Pro Ala Leu Glu Ala Tyr Ser Glu Thr Ala Lys Val Phe Ser Ile Pro
770 775 780

Phe His Lys Asp Cys Gly Glu Asp Gly Leu Cys Ile Ser Asp Leu Val
785 790 795 800

```
Leu Asp Val Arg Gln Ile Pro Ala Ala Gln Glu Gln Pro Phe Ile Val
                805                 810                 815

Ser Asn Gln Asn Lys Arg Leu Thr Phe Ser Val Thr Leu Lys Asn Lys
            820                 825                 830

Arg Glu Ser Ala Tyr Asn Thr Gly Ile Val Val Asp Phe Ser Glu Asn
        835                 840                 845

Leu Phe Phe Ala Ser Phe Ser Leu Pro Val Asp Gly Thr Glu Val Thr
    850                 855                 860

Cys Gln Val Ala Ala Ser Gln Lys Ser Val Ala Cys Asp Val Gly Tyr
865                 870                 875                 880

Pro Ala Leu Lys Arg Glu Gln Gln Val Thr Phe Thr Ile Asn Phe Asp
                885                 890                 895

Phe Asn Leu Gln Asn Leu Gln Asn Gln Ala Ser Leu Ser Phe Gln Ala
            900                 905                 910

Leu Ser Glu Ser Gln Glu Glu Asn Lys Ala Asp Asn Leu Val Asn Leu
        915                 920                 925

Lys Ile Pro Leu Leu Tyr Asp Ala Glu Ile His Leu Thr Arg Ser Thr
    930                 935                 940

Asn Ile Asn Phe Tyr Glu Ile Ser Ser Asp Gly Asn Val Pro Ser Ile
945                 950                 955                 960

Val His Ser Phe Glu Asp Val Gly Pro Lys Phe Ile Phe Ser Leu Lys
                965                 970                 975

Val Thr Thr Gly Ser Val Pro Val Ser Met Ala Thr Val Ile Ile His
            980                 985                 990

Ile Pro Gln Tyr Thr Lys Glu Lys  Asn Pro Leu Met Tyr  Leu Thr Gly
        995                 1000                1005

Val Gln  Thr Asp Lys Ala Gly  Asp Ile Ser Cys Asn  Ala Asp Ile
    1010                1015                1020

Asn Pro  Leu Lys Ile Gly Gln  Thr Ser Ser Ser Val  Ser Phe Lys
    1025                1030                1035
```

```
Ser Glu  Asn Phe Arg His Thr  Lys Glu Leu Asn Cys  Arg Thr Ala
    1040                  1045                  1050

Ser Cys  Ser Asn Val Thr Cys  Trp Leu Lys Asp Val  His Met Lys
    1055                  1060                  1065

Gly Glu  Tyr Phe Val Asn Val  Thr Thr Arg Ile Trp  Asn Gly Thr
    1070                  1075                  1080

Phe Ala  Ser Ser Thr Phe Gln  Thr Val Gln Leu Thr  Ala Ala Ala
    1085                  1090                  1095

Glu Ile  Asn Thr Tyr Asn Pro  Glu Ile Tyr Val Ile  Glu Asp Asn
    1100                  1105                  1110

Thr Val  Thr Ile Pro Leu Met  Ile Met Lys Pro Asp  Glu Lys Ala
    1115                  1120                  1125

Glu Val  Pro Thr Gly Val Ile  Ile Gly Ser Ile Ile  Ala Gly Ile
    1130                  1135                  1140

Leu Leu  Leu Leu Ala Leu Val  Ala Ile Leu Trp Lys  Leu Gly Phe
    1145                  1150                  1155

Phe Lys  Arg Lys Tyr Glu Lys  Met Thr Lys Asn Pro  Asp Glu Ile
    1160                  1165                  1170

Asp Glu  Thr Thr Glu Leu Ser  Ser
    1175                  1180
```

<210> 78
<211> 332
<212> PRT
<213> Homo Sapiens

<400> 78

```
Met Tyr Arg Pro Arg Ala Arg Ala Ala Pro Glu Gly Arg Val Arg Gly
1               5                  10                  15

Cys Ala Val Pro Ser Thr Val Leu Leu Leu Leu Ala Tyr Leu Ala Tyr
            20                  25                  30

Leu Ala Leu Gly Thr Gly Val Phe Trp Thr Leu Glu Gly Arg Ala Ala
        35                  40                  45

Gln Asp Ser Ser Arg Ser Phe Gln Arg Asp Lys Trp Glu Leu Leu Gln
```

50 55 60

Asn Phe Thr Cys Leu Asp Arg Pro Ala Leu Asp Ser Leu Ile Arg Asp
65 70 75 80

Val Val Gln Ala Tyr Lys Asn Gly Ala Ser Leu Leu Ser Asn Thr Thr
85 90 95

Ser Met Gly Arg Trp Glu Leu Val Gly Ser Phe Phe Phe Ser Val Ser
100 105 110

Thr Ile Thr Thr Ile Gly Tyr Gly Asn Leu Ser Pro Asn Thr Met Ala
115 120 125

Ala Arg Leu Phe Cys Ile Phe Phe Ala Leu Val Gly Ile Pro Leu Asn
130 135 140

Leu Val Val Leu Asn Arg Leu Gly His Leu Met Gln Gln Gly Val Asn
145 150 155 160

His Trp Ala Ser Arg Leu Gly Gly Thr Trp Gln Asp Pro Asp Lys Ala
165 170 175

Arg Trp Leu Ala Gly Ser Gly Ala Leu Leu Ser Gly Leu Leu Leu Phe
180 185 190

Leu Leu Leu Pro Pro Leu Leu Phe Ser His Met Glu Gly Trp Ser Tyr
195 200 205

Thr Glu Gly Phe Tyr Phe Ala Phe Ile Thr Leu Ser Thr Val Gly Phe
210 215 220

Gly Asp Tyr Val Ile Gly Met Asn Pro Ser Gln Arg Tyr Pro Leu Trp
225 230 235 240

Tyr Lys Asn Met Val Ser Leu Trp Ile Leu Phe Gly Met Ala Trp Leu
245 250 255

Ala Leu Ile Ile Lys Leu Ile Leu Ser Gln Leu Glu Thr Pro Gly Arg
260 265 270

Val Cys Ser Cys Cys His His Ser Ser Lys Glu Asp Phe Lys Ser Gln
275 280 285

Ser Trp Arg Gln Gly Pro Asp Arg Glu Pro Glu Ser His Ser Pro Gln
290 295 300

Gln Gly Cys Tyr Pro Glu Gly Pro Met Gly Ile Ile Gln His Leu Glu
305 310 315 320

Pro Ser Ala His Ala Ala Gly Cys Gly Lys Asp Ser
325 330

<210> 79
<211> 328
<212> PRT
<213> Homo Sapiens

<400> 79

Met Glu Trp Asp Asn Gly Thr Gly Gln Ala Leu Gly Leu Pro Pro Thr
1 5 10 15

Thr Cys Val Tyr Arg Glu Asn Phe Lys Gln Leu Leu Leu Pro Pro Val
20 25 30

Tyr Ser Ala Val Leu Ala Ala Gly Leu Pro Leu Asn Ile Cys Val Ile
35 40 45

Thr Gln Ile Cys Thr Ser Arg Arg Ala Leu Thr Arg Thr Ala Val Tyr
50 55 60

Thr Leu Asn Leu Ala Leu Ala Asp Leu Leu Tyr Ala Cys Ser Leu Pro
65 70 75 80

Leu Leu Ile Tyr Asn Tyr Ala Gln Gly Asp His Trp Pro Phe Gly Asp
85 90 95

Phe Ala Cys Arg Leu Val Arg Phe Leu Phe Tyr Ala Asn Leu His Gly
100 105 110

Ser Ile Leu Phe Leu Thr Cys Ile Ser Phe Gln Arg Tyr Leu Gly Ile
115 120 125

Cys His Pro Leu Ala Pro Trp His Lys Arg Gly Gly Arg Arg Ala Ala
130 135 140

Trp Leu Val Cys Val Ala Val Trp Leu Ala Val Thr Thr Gln Cys Leu
145 150 155 160

Pro Thr Ala Ile Phe Ala Ala Thr Gly Ile Gln Arg Asn Arg Thr Val
165 170 175

Cys Tyr Asp Leu Ser Pro Pro Ala Leu Ala Thr His Tyr Met Pro Tyr
180 185 190

Gly Met Ala Leu Thr Val Ile Gly Phe Leu Leu Pro Phe Ala Ala Leu
195 200 205

Leu Ala Cys Tyr Cys Leu Leu Ala Cys Arg Leu Cys Arg Gln Asp Gly
210 215 220

Pro Ala Glu Pro Val Ala Gln Glu Arg Arg Gly Lys Ala Ala Arg Met
225 230 235 240

Ala Val Val Val Ala Ala Ala Phe Ala Ile Ser Phe Leu Pro Phe His
245 250 255

Ile Thr Lys Thr Ala Tyr Leu Ala Val Arg Ser Thr Pro Gly Val Pro
260 265 270

Cys Thr Val Leu Glu Ala Phe Ala Ala Ala Tyr Lys Gly Thr Arg Pro
275 280 285

Phe Ala Ser Ala Asn Ser Val Leu Asp Pro Ile Leu Phe Tyr Phe Thr
290 295 300

Gln Lys Lys Phe Arg Arg Arg Pro His Glu Leu Leu Gln Lys Leu Thr
305 310 315 320

Ala Lys Trp Gln Arg Gln Gly Arg
325

<210> 80
<211> 581
<212> PRT
<213> Homo Sapiens

<400> 80

Met Gln Arg Pro Gly Pro Arg Leu Trp Leu Val Leu Gln Val Met Gly
1 5 10 15

Ser Cys Ala Ala Ile Ser Ser Met Asp Met Glu Arg Pro Gly Asp Gly
20 25 30

Lys Cys Gln Pro Ile Glu Ile Pro Met Cys Lys Asp Ile Gly Tyr Asn
35 40 45

Met Thr Arg Met Pro Asn Leu Met Gly His Glu Asn Gln Arg Glu Ala
50 55 60

Ala Ile Gln Leu His Glu Phe Ala Pro Leu Val Glu Tyr Gly Cys His
65 70 75 80

Gly His Leu Arg Phe Phe Leu Cys Ser Leu Tyr Ala Pro Met Cys Thr
85 90 95

Glu Gln Val Ser Thr Pro Ile Pro Ala Cys Arg Val Met Cys Glu Gln
100 105 110

Ala Arg Leu Lys Cys Ser Pro Ile Met Glu Gln Phe Asn Phe Lys Trp
115 120 125

Pro Asp Ser Leu Asp Cys Arg Lys Leu Pro Asn Lys Asn Asp Pro Asn
130 135 140

Tyr Leu Cys Met Glu Ala Pro Asn Asn Gly Ser Asp Glu Pro Thr Arg
145 150 155 160

Gly Ser Gly Leu Phe Pro Pro Leu Phe Arg Pro Gln Arg Pro His Ser
165 170 175

Ala Gln Glu His Pro Leu Lys Asp Gly Gly Pro Gly Arg Gly Gly Cys
180 185 190

Asp Asn Pro Gly Lys Phe His His Val Glu Lys Ser Ala Ser Cys Ala
195 200 205

Pro Leu Cys Thr Pro Gly Val Asp Val Tyr Trp Ser Arg Glu Asp Lys
210 215 220

Arg Phe Ala Val Val Trp Leu Ala Ile Trp Ala Val Leu Cys Phe Phe
225 230 235 240

Ser Ser Ala Phe Thr Val Leu Thr Phe Leu Ile Asp Pro Ala Arg Phe
245 250 255

Arg Tyr Pro Glu Arg Pro Ile Ile Phe Leu Ser Met Cys Tyr Cys Val
260 265 270

Tyr Ser Val Gly Tyr Leu Ile Arg Leu Phe Ala Gly Ala Glu Ser Ile
275 280 285

Ala Cys Asp Arg Asp Ser Gly Gln Leu Tyr Val Ile Gln Glu Gly Leu
290 295 300

Glu Ser Thr Gly Cys Thr Leu Val Phe Leu Val Leu Tyr Tyr Phe Gly
305 310 315 320

Met Ala Ser Ser Leu Trp Trp Val Val Leu Thr Leu Thr Trp Phe Leu
325 330 335

Ala Ala Gly Lys Lys Trp Gly His Glu Ala Ile Glu Ala Asn Ser Ser
340 345 350

Tyr Phe His Leu Ala Ala Trp Ala Ile Pro Ala Val Lys Thr Ile Leu
355 360 365

Ile Leu Val Met Arg Arg Val Ala Gly Asp Glu Leu Thr Gly Val Cys
370 375 380

Tyr Val Gly Ser Met Asp Val Asn Ala Leu Thr Gly Phe Val Leu Ile
385 390 395 400

Pro Leu Ala Cys Tyr Leu Val Ile Gly Thr Ser Phe Ile Leu Ser Gly
405 410 415

Phe Val Ala Leu Phe His Ile Arg Arg Val Met Lys Thr Gly Gly Glu
420 425 430

Asn Thr Asp Lys Leu Glu Lys Leu Met Val Arg Ile Gly Leu Phe Ser
435 440 445

Val Leu Tyr Thr Val Pro Ala Thr Cys Val Ile Ala Cys Tyr Phe Tyr
450 455 460

Glu Arg Leu Asn Met Asp Tyr Trp Lys Ile Leu Ala Ala Gln His Lys
465 470 475 480

Cys Lys Met Asn Asn Gln Thr Lys Thr Leu Asp Cys Leu Met Ala Ala
485 490 495

Ser Ile Pro Ala Val Glu Ile Phe Met Val Lys Ile Phe Met Leu Leu
500 505 510

Val Val Gly Ile Thr Ser Gly Met Trp Ile Trp Thr Ser Lys Thr Leu
515 520 525

Gln Ser Trp Gln Gln Val Cys Ser Arg Arg Leu Lys Lys Lys Ser Arg
530 535 540

Arg Lys Pro Ala Ser Val Ile Thr Ser Gly Gly Ile Tyr Lys Lys Ala
545 550 555 560

Gln His Pro Gln Lys Thr His His Gly Lys Tyr Glu Ile Pro Ala Gln
565 570 575

Ser Pro Thr Cys Val
580

<210> 81
<211> 539
<212> PRT
<213> Homo sapiens

<400> 81

Met Val Pro Gly Ala Arg Gly Gly Gly Ala Leu Ala Arg Ala Ala Gly
1 5 10 15

Arg Gly Leu Leu Ala Leu Leu Leu Ala Val Ser Ala Pro Leu Arg Leu
20 25 30

Gln Ala Glu Glu Leu Gly Asp Gly Cys Gly His Leu Val Thr Tyr Gln
35 40 45

Asp Ser Gly Thr Met Thr Ser Lys Asn Tyr Pro Gly Thr Tyr Pro Asn
50 55 60

His Thr Val Cys Glu Lys Thr Ile Thr Val Pro Lys Gly Lys Arg Leu
65 70 75 80

Ile Leu Arg Leu Gly Asp Leu Asp Ile Glu Ser Gln Thr Cys Ala Ser
85 90 95

Asp Tyr Leu Leu Phe Thr Ser Ser Ser Asp Gln Tyr Gly Pro Tyr Cys
100 105 110

Gly Ser Met Thr Val Pro Lys Glu Leu Leu Leu Asn Thr Ser Glu Val
115 120 125

Thr Val Arg Phe Glu Ser Gly Ser His Ile Ser Gly Arg Gly Phe Leu
130 135 140

Leu Thr Tyr Ala Ser Ser Asp His Pro Asp Leu Ile Thr Cys Leu Glu
145 150 155 160

Arg Ala Ser His Tyr Leu Lys Thr Glu Tyr Ser Lys Phe Cys Pro Ala
165 170 175

Gly Cys Arg Asp Val Ala Gly Asp Ile Ser Gly Asn Met Val Asp Gly
180 185 190

Tyr Arg Asp Thr Ser Leu Leu Cys Lys Ala Ala Ile His Ala Gly Ile
195 200 205

Ile Ala Asp Glu Leu Gly Gly Gln Ile Ser Val Leu Gln Arg Lys Gly
210 215 220

Ile Ser Arg Tyr Glu Gly Ile Leu Ala Asn Gly Val Leu Ser Arg Asp
225 230 235 240

Gly Ser Leu Ser Asp Lys Arg Phe Leu Phe Thr Ser Asn Gly Cys Ser
245 250 255

Arg Ser Leu Ser Phe Glu Pro Asp Gly Gln Ile Arg Ala Ser Ser Ser
260 265 270

Trp Gln Ser Val Asn Glu Ser Gly Asp Gln Val His Trp Ser Pro Gly
275 280 285

Gln Ala Arg Leu Gln Asp Gln Gly Pro Ser Trp Ala Ser Gly Asp Ser
290 295 300

Ser Asn Asn His Lys Pro Arg Glu Trp Leu Glu Ile Asp Leu Gly Glu
305 310 315 320

Lys Lys Lys Ile Thr Gly Ile Arg Thr Thr Gly Ser Thr Gln Ser Asn
325 330 335

Phe Asn Phe Tyr Val Lys Ser Phe Val Met Asn Phe Lys Asn Asn Asn
340 345 350

Ser Lys Trp Lys Thr Tyr Lys Gly Ile Val Asn Asn Glu Glu Lys Val
355 360 365

Phe Gln Gly Asn Ser Asn Phe Arg Asp Pro Val Gln Asn Asn Phe Ile
370 375 380

Pro Pro Ile Val Ala Arg Tyr Val Arg Val Val Pro Gln Thr Trp His
385 390 395 400

Gln Arg Ile Ala Leu Lys Val Glu Leu Ile Gly Cys Gln Ile Thr Gln
405 410 415

Gly Asn Asp Ser Leu Val Trp Arg Lys Thr Ser Gln Ser Thr Ser Val
420 425 430

Ser Thr Lys Lys Glu Asp Glu Thr Ile Thr Arg Pro Ile Pro Ser Glu
435 440 445

Glu Thr Ser Thr Gly Ile Asn Ile Thr Thr Val Ala Ile Pro Leu Val
450 455 460

Leu Leu Val Val Leu Val Phe Ala Gly Met Gly Ile Phe Ala Ala Phe
465 470 475 480

Arg Lys Lys Lys Lys Lys Gly Ser Pro Tyr Gly Ser Ala Glu Ala Gln
485 490 495

Lys Thr Asp Cys Trp Lys Gln Ile Lys Tyr Pro Phe Ala Arg His Gln
500 505 510

Ser Ala Glu Phe Thr Ile Ser Tyr Asp Asn Glu Lys Glu Met Thr Gln
515 520 525

Lys Leu Asp Leu Ile Thr Ser Asp Met Ala Gly
530 535

<210> 82
<211> 539
<212> PRT
<213> Homo sapiens

<400> 82

Met Val Pro Gly Ala Arg Gly Gly Gly Ala Leu Ala Arg Ala Ala Gly
1 5 10 15

Arg Gly Leu Leu Ala Leu Leu Leu Ala Val Ser Ala Pro Leu Arg Leu
20 25 30

Gln Ala Glu Glu Leu Gly Asp Gly Cys Gly His Leu Val Thr Tyr Gln
35 40 45

Asp Ser Gly Thr Met Thr Ser Lys Asn Tyr Pro Gly Thr Tyr Pro Asn
50 55 60

His Thr Val Cys Glu Lys Thr Ile Thr Val Pro Lys Gly Lys Arg Leu
65 70 75 80

Ile Leu Arg Leu Gly Asp Leu Asp Ile Glu Ser Gln Thr Cys Ala Ser

85 90 95

Asp Tyr Leu Leu Phe Thr Ser Ser Ser Asp Gln Tyr Gly Pro Tyr Cys
100 105 110

Gly Ser Met Thr Val Pro Lys Glu Leu Leu Leu Asn Thr Ser Glu Val
115 120 125

Thr Val Arg Phe Glu Ser Gly Ser His Ile Ser Gly Arg Gly Phe Leu
130 135 140

Leu Thr Tyr Ala Ser Ser Asp His Pro Asp Leu Ile Thr Cys Leu Glu
145 150 155 160

Arg Ala Ser His Tyr Leu Lys Thr Glu Tyr Ser Lys Phe Cys Pro Ala
165 170 175

Gly Cys Arg Asp Val Ala Gly Asp Ile Ser Gly Asn Met Val Asp Gly
180 185 190

Tyr Arg Asp Thr Ser Leu Leu Cys Lys Ala Ala Ile His Ala Gly Ile
195 200 205

Ile Ala Asp Glu Leu Gly Gly Gln Ile Ser Val Leu Gln Arg Lys Gly
210 215 220

Ile Ser Arg Tyr Glu Gly Ile Leu Ala Asn Gly Val Leu Ser Arg Asp
225 230 235 240

Gly Ser Leu Ser Asp Lys Arg Phe Leu Phe Thr Ser Asn Gly Cys Ser
245 250 255

Arg Ser Leu Ser Phe Glu Pro Asp Gly Gln Ile Arg Ala Ser Ser Ser
260 265 270

Trp Gln Ser Val Asn Glu Ser Gly Asp Gln Val His Trp Ser Pro Gly
275 280 285

Gln Ala Arg Leu Gln Asp Gln Gly Pro Ser Trp Ala Ser Gly Asp Ser
290 295 300

Ser Asn Asn His Lys Pro Arg Glu Trp Leu Glu Ile Asp Leu Gly Glu
305 310 315 320

Lys Lys Lys Ile Thr Gly Ile Arg Thr Thr Gly Ser Thr Gln Ser Asn
325 330 335

Phe Asn Phe Tyr Val Lys Ser Phe Val Met Asn Phe Lys Asn Asn Asn
340 345 350

Ser Lys Trp Lys Thr Tyr Lys Gly Ile Val Asn Asn Glu Glu Lys Val
355 360 365

Phe Gln Gly Asn Ser Asn Phe Arg Asp Pro Val Gln Asn Asn Phe Ile
370 375 380

Pro Pro Ile Val Ala Arg Tyr Val Arg Val Val Pro Gln Thr Trp His
385 390 395 400

Gln Arg Ile Ala Leu Lys Val Glu Leu Ile Gly Cys Gln Ile Thr Gln
405 410 415

Gly Asn Asp Ser Leu Val Trp Arg Lys Thr Ser Gln Ser Thr Ser Val
420 425 430

Ser Thr Lys Lys Glu Asp Glu Thr Ile Thr Arg Pro Ile Pro Ser Glu
435 440 445

Glu Thr Ser Thr Gly Ile Asn Ile Thr Thr Val Ala Ile Pro Leu Val
450 455 460

Leu Leu Val Val Leu Val Phe Ala Gly Met Gly Ile Phe Ala Ala Phe
465 470 475 480

Arg Lys Lys Lys Lys Lys Gly Ser Pro Tyr Gly Ser Ala Glu Ala Gln
485 490 495

Lys Thr Asp Cys Trp Lys Gln Ile Lys Tyr Pro Phe Ala Arg His Gln
500 505 510

Ser Ala Glu Phe Thr Ile Ser Tyr Asp Asn Glu Lys Glu Met Thr Gln
515 520 525

Lys Leu Asp Leu Ile Thr Ser Asp Met Ala Gly
530 535

<210> 83
<211> 237
<212> PRT
<213> Homo Sapiens

<400> 83

Met Ala Gly Val Ser Ala Cys Ile Lys Tyr Ser Met Phe Thr Phe Asn
1 5 10 15

Phe Leu Phe Trp Leu Cys Gly Ile Leu Ile Leu Ala Leu Ala Ile Trp
20 25 30

Val Arg Val Ser Asn Asp Ser Gln Ala Ile Phe Gly Ser Glu Asp Val
35 40 45

Gly Ser Ser Ser Tyr Val Ala Val Asp Ile Leu Ile Ala Val Gly Ala
50 55 60

Ile Ile Met Ile Leu Gly Phe Leu Gly Cys Cys Gly Ala Ile Lys Glu
65 70 75 80

Ser Arg Cys Met Leu Leu Leu Phe Phe Ile Gly Leu Leu Leu Ile Leu
85 90 95

Leu Leu Gln Val Ala Thr Gly Ile Leu Gly Ala Val Phe Lys Ser Lys
100 105 110

Ser Asp Arg Ile Val Asn Glu Thr Leu Tyr Glu Asn Thr Lys Leu Leu
115 120 125

Ser Ala Thr Gly Glu Ser Glu Lys Gln Phe Gln Glu Ala Ile Ile Val
130 135 140

Phe Gln Glu Glu Phe Lys Cys Cys Gly Leu Val Asn Gly Ala Ala Asp
145 150 155 160

Trp Gly Asn Asn Phe Gln His Tyr Pro Glu Leu Cys Ala Cys Leu Asp
165 170 175

Lys Gln Arg Pro Cys Gln Ser Tyr Asn Gly Lys Gln Val Tyr Lys Glu
180 185 190

Thr Cys Ile Ser Phe Ile Lys Asp Phe Leu Ala Lys Asn Leu Ile Ile
195 200 205

Val Ile Gly Ile Ser Phe Gly Leu Ala Val Ile Glu Ile Leu Gly Leu
210 215 220

Val Phe Ser Met Val Leu Tyr Cys Gln Ile Gly Asn Lys
225 230 235

<210> 84
<211> 202
<212> PRT
<213> Homo Sapiens

<400> 84

Met Cys Thr Gly Gly Cys Ala Arg Cys Leu Gly Gly Thr Leu Ile Pro
1 5 10 15

Leu Ala Phe Phe Gly Phe Leu Ala Asn Ile Leu Leu Phe Phe Pro Gly
20 25 30

Gly Lys Val Ile Asp Asp Asn Asp His Leu Ser Gln Glu Ile Trp Phe
35 40 45

Phe Gly Gly Ile Leu Gly Ser Gly Val Leu Met Ile Phe Pro Ala Leu
50 55 60

Val Phe Leu Gly Leu Lys Asn Asn Asp Cys Cys Gly Cys Cys Gly Asn
65 70 75 80

Glu Gly Cys Gly Lys Arg Phe Ala Met Phe Thr Ser Thr Ile Phe Ala
85 90 95

Val Val Gly Phe Leu Gly Ala Gly Tyr Ser Phe Ile Ile Ser Ala Ile
100 105 110

Ser Ile Asn Lys Gly Pro Lys Cys Leu Met Ala Asn Ser Thr Trp Gly
115 120 125

Tyr Pro Phe His Asp Gly Asp Tyr Leu Asn Asp Glu Ala Leu Trp Asn
130 135 140

Lys Cys Arg Glu Pro Leu Asn Val Val Pro Trp Asn Leu Thr Leu Phe
145 150 155 160

Ser Ile Leu Leu Val Val Gly Gly Ile Gln Met Val Leu Cys Ala Ile
165 170 175

Gln Val Val Asn Gly Leu Leu Gly Thr Leu Cys Gly Asp Cys Gln Cys
180 185 190

Cys Gly Cys Cys Gly Gly Asp Gly Pro Val
195 200

<210> 85
<211> 677

<212> PRT
<213> Homo Sapiens

<400> 85

Met Gln Pro Thr Leu Leu Leu Ser Leu Leu Gly Ala Val Gly Leu Ala
1 5 10 15

Ala Val Asn Ser Met Pro Val Asp Asn Arg Asn His Asn Glu Gly Met
20 25 30

Val Thr Arg Cys Ile Ile Glu Val Leu Ser Asn Ala Leu Ser Lys Ser
35 40 45

Ser Ala Pro Pro Ile Thr Pro Glu Cys Arg Gln Val Leu Lys Thr Ser
50 55 60

Arg Lys Asp Val Lys Asp Lys Glu Thr Thr Glu Asn Glu Asn Thr Lys
65 70 75 80

Phe Glu Val Arg Leu Leu Arg Asp Pro Ala Asp Ala Ser Glu Ala His
85 90 95

Glu Ser Ser Ser Arg Gly Glu Ala Gly Ala Pro Gly Glu Glu Asp Ile
100 105 110

Gln Gly Pro Thr Lys Ala Asp Thr Glu Lys Trp Ala Glu Gly Gly Gly
115 120 125

His Ser Arg Glu Arg Ala Asp Glu Pro Gln Trp Ser Leu Tyr Pro Ser
130 135 140

Asp Ser Gln Val Ser Glu Glu Val Lys Thr Arg His Ser Glu Lys Ser
145 150 155 160

Gln Arg Glu Asp Glu Glu Glu Glu Glu Gly Glu Asn Tyr Gln Lys Gly
165 170 175

Glu Arg Gly Glu Asp Ser Ser Glu Glu Lys His Leu Glu Glu Pro Gly
180 185 190

Glu Thr Gln Asn Ala Phe Leu Asn Glu Arg Lys Gln Ala Ser Ala Ile
195 200 205

Lys Lys Glu Glu Leu Val Ala Arg Ser Glu Thr His Ala Ala Gly His
210 215 220

Ser Gln Glu Lys Thr His Ser Arg Glu Lys Ser Ser Gln Glu Ser Gly
225 230 235 240

Glu Glu Ala Gly Ser Gln Glu Asn His Pro Gln Glu Ser Lys Gly Gln
245 250 255

Pro Arg Ser Gln Glu Glu Ser Glu Glu Gly Glu Glu Asp Ala Thr Ser
260 265 270

Glu Val Asp Lys Arg Arg Thr Arg Pro Arg His His His Gly Arg Ser
275 280 285

Arg Pro Asp Arg Ser Ser Gln Gly Gly Ser Leu Pro Ser Glu Glu Lys
290 295 300

Gly His Pro Gln Glu Glu Ser Glu Glu Ser Asn Val Ser Met Ala Ser
305 310 315 320

Leu Gly Glu Lys Arg Asp His His Ser Thr His Tyr Arg Ala Ser Glu
325 330 335

Glu Glu Pro Glu Tyr Gly Glu Glu Ile Lys Gly Tyr Pro Gly Val Gln
340 345 350

Ala Pro Glu Asp Leu Glu Trp Glu Arg Tyr Arg Gly Arg Gly Ser Glu
355 360 365

Glu Tyr Arg Ala Pro Arg Pro Gln Ser Glu Glu Ser Trp Asp Glu Glu
370 375 380

Asp Lys Arg Asn Tyr Pro Ser Leu Glu Leu Asp Lys Met Ala His Gly
385 390 395 400

Tyr Gly Glu Glu Ser Glu Glu Glu Arg Gly Leu Glu Pro Gly Lys Gly
405 410 415

Arg His His Arg Gly Arg Gly Gly Glu Pro Arg Ala Tyr Phe Met Ser
420 425 430

Asp Thr Arg Glu Glu Lys Arg Phe Leu Gly Glu Gly His His Arg Val
435 440 445

Gln Glu Asn Gln Met Asp Lys Ala Arg Arg His Pro Gln Gly Ala Trp
450 455 460

Lys Glu Leu Asp Arg Asn Tyr Leu Asn Tyr Gly Glu Glu Gly Ala Pro

465 470 475 480

Gly Lys Trp Gln Gln Gln Gly Asp Leu Gln Asp Thr Lys Glu Asn Arg
485 490 495

Glu Glu Ala Arg Phe Gln Asp Lys Gln Tyr Ser Ser His His Thr Ala
500 505 510

Glu Lys Arg Lys Arg Leu Gly Glu Leu Phe Asn Pro Tyr Tyr Asp Pro
515 520 525

Leu Gln Trp Lys Ser Ser His Phe Glu Arg Arg Asp Asn Met Asn Asp
530 535 540

Asn Phe Leu Glu Gly Glu Glu Glu Asn Glu Leu Thr Leu Asn Glu Lys
545 550 555 560

Asn Phe Phe Pro Glu Tyr Asn Tyr Asp Trp Trp Glu Lys Lys Pro Phe
565 570 575

Ser Glu Asp Val Asn Trp Gly Tyr Glu Lys Arg Asn Leu Ala Arg Val
580 585 590

Pro Lys Leu Asp Leu Lys Arg Gln Tyr Asp Arg Val Ala Gln Leu Asp
595 600 605

Gln Leu Leu His Tyr Arg Lys Lys Ser Ala Glu Phe Pro Asp Phe Tyr
610 615 620

Asp Ser Glu Glu Pro Val Ser Thr His Gln Glu Ala Glu Asn Glu Lys
625 630 635 640

Asp Arg Ala Asp Gln Thr Val Leu Thr Glu Asp Glu Lys Lys Glu Leu
645 650 655

Glu Asn Leu Ala Ala Met Asp Leu Glu Leu Gln Lys Ile Ala Glu Lys
660 665 670

Phe Ser Gln Arg Gly
675

<210> 86
<211> 631
<212> PRT
<213> Homo Sapiens

<400> 86

Met Lys Leu Leu Cys Glu Gly Leu Lys Gln Pro Asn Cys Val Leu Gln
1 5 10 15

Thr Leu Arg Trp Tyr Arg Cys Leu Ile Ser Ser Ala Ser Cys Gly Ala
20 25 30

Leu Ala Ala Val Leu Ser Thr Ser Gln Trp Leu Thr Glu Leu Glu Phe
35 40 45

Ser Glu Thr Lys Leu Glu Ala Ser Ala Leu Lys Leu Leu Tyr Gly Gly
50 55 60

Leu Lys Asp Pro Asn Cys Lys Leu Gln Lys Leu Asn Leu Gln Phe Ser
65 70 75 80

Leu Ser Val Thr Ala Ala Lys Leu Pro Val Gly Met Val Gly Asn Cys
85 90 95

Ser Gly Phe Ser Gly Ser Leu Val Gln Ser His Phe Gly Tyr Cys Gln
100 105 110

Asp Ser Ser Phe Lys Cys Asp Leu Cys Lys Leu Leu Trp Pro Ser Thr
115 120 125

Arg Val Ala Ala Ala Lys Asp Cys Gly Ser Pro Lys Ser Phe Leu Ser
130 135 140

Glu Gly Leu Asn Trp Ala Gly Arg Leu Glu Ala Val Glu Glu Val Leu
145 150 155 160

Gly Leu Gly Val Leu Val Gln Pro Gly Asp Pro Ala Ser Gln Gly Gly
165 170 175

Gly His Cys Glu Asn Tyr Gly Ser Phe Arg Asp Leu Val Asp Leu Glu
180 185 190

Val Lys Ala Glu Pro Ser Leu Arg Lys Gly Gly Met Asp Leu Gln Arg
195 200 205

Pro Thr Leu Gln Val Val Leu Leu Cys Lys Ile Phe Ser Leu Lys Leu
210 215 220

Phe Leu Phe Ile Ala Leu Pro Asn Ser Pro Gly Gln Val Ser Val Val
225 230 235 240

```
Gln Val Thr Ile Pro Asp Gly Phe Val Asn Val Thr Val Gly Ser Asn
                245                 250                 255

Val Thr Leu Ile Cys Ile Tyr Thr Thr Thr Val Ala Ser Arg Glu Gln
            260                 265                 270

Leu Ser Ile Gln Trp Ser Phe Phe His Lys Lys Glu Met Glu Pro Ile
        275                 280                 285

Ser Ser Pro Trp Glu Glu Gly Lys Trp Pro Asp Val Glu Ala Val Lys
    290                 295                 300

Gly Thr Leu Asp Gly Gln Gln Ala Glu Leu Gln Ile Tyr Phe Ser Gln
305                 310                 315                 320

Gly Gly Gln Ala Val Ala Ile Gly Gln Phe Lys Asp Arg Ile Thr Gly
                325                 330                 335

Ser Asn Asp Pro Gly Asn Ala Ser Ile Thr Ile Ser His Met Gln Pro
            340                 345                 350

Ala Asp Ser Gly Ile Tyr Ile Cys Asp Val Asn Asn Pro Pro Asp Phe
        355                 360                 365

Leu Gly Gln Asn Gln Gly Ile Leu Asn Val Ser Val Leu Val Lys Pro
    370                 375                 380

Ser Lys Pro Leu Cys Ser Val Gln Gly Arg Pro Glu Thr Gly His Thr
385                 390                 395                 400

Ile Ser Leu Ser Cys Leu Ser Ala Leu Gly Thr Pro Ser Pro Val Tyr
                405                 410                 415

Tyr Trp His Lys Leu Glu Gly Arg Asp Ile Val Pro Val Lys Glu Asn
            420                 425                 430

Phe Asn Pro Thr Thr Gly Ile Leu Val Ile Gly Asn Leu Thr Asn Phe
        435                 440                 445

Glu Gln Gly Tyr Tyr Gln Cys Thr Ala Ile Asn Arg Leu Gly Asn Ser
    450                 455                 460

Ser Cys Glu Ile Asp Leu Thr Ser Ser His Pro Glu Val Gly Ile Ile
465                 470                 475                 480

Val Gly Ala Leu Ile Gly Ser Leu Val Gly Ala Ala Ile Ile Ile Ser
```

485 490 495

Val Val Cys Phe Ala Arg Asn Lys Ala Lys Ala Lys Ala Lys Glu Arg
500 505 510

Asn Ser Lys Thr Ile Ala Glu Leu Glu Pro Met Thr Lys Ile Asn Pro
515 520 525

Arg Gly Glu Ser Glu Ala Met Pro Arg Glu Asp Ala Thr Gln Leu Glu
530 535 540

Val Thr Leu Pro Ser Ser Ile His Glu Thr Gly Pro Asp Thr Ile Gln
545 550 555 560

Glu Pro Asp Tyr Glu Pro Lys Pro Thr Gln Glu Pro Ala Pro Glu Pro
565 570 575

Ala Pro Gly Ser Glu Pro Met Ala Val Pro Asp Leu Asp Ile Glu Leu
580 585 590

Glu Leu Glu Pro Glu Thr Gln Ser Glu Leu Glu Pro Glu Pro Glu Pro
595 600 605

Glu Pro Glu Ser Glu Pro Gly Val Val Val Glu Pro Leu Ser Glu Asp
610 615 620

Glu Lys Gly Val Val Lys Ala
625 630

<210> 87
<211> 413
<212> PRT
<213> Homo Sapiens

<400> 87

Met Val Phe Ala Phe Trp Lys Val Phe Leu Ile Leu Ser Cys Leu Ala
1 5 10 15

Gly Gln Val Ser Val Val Gln Val Thr Ile Pro Asp Gly Phe Val Asn
20 25 30

Val Thr Val Gly Ser Asn Val Thr Leu Ile Cys Ile Tyr Thr Thr Thr
35 40 45

Val Ala Ser Arg Glu Gln Leu Ser Ile Gln Trp Ser Phe Phe His Lys
50 55 60

Lys Glu Met Glu Pro Ile Ser Ser Pro Trp Glu Glu Gly Lys Trp Pro
65 70 75 80

Asp Val Glu Ala Val Lys Gly Thr Leu Asp Gly Gln Gln Ala Glu Leu
85 90 95

Gln Ile Tyr Phe Ser Gln Gly Gly Gln Ala Val Ala Ile Gly Gln Phe
100 105 110

Lys Asp Arg Ile Thr Gly Ser Asn Asp Pro Gly Asn Ala Ser Ile Thr
115 120 125

Ile Ser His Met Gln Pro Ala Asp Ser Gly Ile Tyr Ile Cys Asp Val
130 135 140

Asn Asn Pro Pro Asp Phe Leu Gly Gln Asn Gln Gly Ile Leu Asn Val
145 150 155 160

Ser Val Leu Val Lys Pro Ser Lys Pro Leu Cys Ser Val Gln Gly Arg
165 170 175

Pro Glu Thr Gly His Thr Ile Ser Leu Ser Cys Leu Ser Ala Leu Gly
180 185 190

Thr Pro Ser Pro Val Tyr Tyr Trp His Lys Leu Glu Gly Arg Asp Ile
195 200 205

Val Pro Val Lys Glu Asn Phe Asn Pro Thr Thr Gly Ile Leu Val Ile
210 215 220

Gly Asn Leu Thr Asn Phe Glu Gln Gly Tyr Tyr Gln Cys Thr Ala Ile
225 230 235 240

Asn Arg Leu Gly Asn Ser Ser Cys Glu Ile Asp Leu Thr Ser Ser His
245 250 255

Pro Glu Val Gly Ile Ile Val Gly Ala Leu Ile Gly Ser Leu Val Gly
260 265 270

Ala Ala Ile Ile Ile Ser Val Val Cys Phe Ala Arg Asn Lys Ala Lys
275 280 285

Ala Lys Ala Lys Glu Arg Asn Ser Lys Thr Ile Ala Glu Leu Glu Pro
290 295 300

```
Met Thr Lys Ile Asn Pro Arg Gly Glu Ser Glu Ala Met Pro Arg Glu
305                 310                 315                 320

Asp Ala Thr Gln Leu Glu Val Thr Leu Pro Ser Ser Ile His Glu Thr
                325                 330                 335

Gly Pro Asp Thr Ile Gln Glu Pro Asp Tyr Glu Pro Lys Pro Thr Gln
            340                 345                 350

Glu Pro Ala Pro Glu Pro Ala Pro Gly Ser Glu Pro Met Ala Val Pro
        355                 360                 365

Asp Leu Asp Ile Glu Leu Glu Leu Glu Pro Glu Thr Gln Ser Glu Leu
    370                 375                 380

Glu Pro Glu Pro Glu Pro Glu Pro Glu Ser Glu Pro Gly Val Val Val
385                 390                 395                 400

Glu Pro Leu Ser Glu Asp Glu Lys Gly Val Val Lys Ala
                405                 410
```

<210> 88
<211> 397
<212> PRT
<213> Homo Sapiens

<400> 88

```
Met Arg Ser Pro Ser Ala Ala Trp Leu Leu Gly Ala Ala Ile Leu Leu
1               5                   10                  15

Ala Ala Ser Leu Ser Cys Ser Gly Thr Ile Gln Gly Thr Asn Arg Ser
            20                  25                  30

Ser Lys Gly Arg Ser Leu Ile Gly Lys Val Asp Gly Thr Ser His Val
        35                  40                  45

Thr Gly Lys Gly Val Thr Val Glu Thr Val Phe Ser Val Asp Glu Phe
    50                  55                  60

Ser Ala Ser Val Leu Thr Gly Lys Leu Thr Thr Val Phe Leu Pro Ile
65                  70                  75                  80

Val Tyr Thr Ile Val Phe Val Val Gly Leu Pro Ser Asn Gly Met Ala
                85                  90                  95

Leu Trp Val Phe Leu Phe Arg Thr Lys Lys Lys His Pro Ala Val Ile
            100                 105                 110
```

Tyr Met Ala Asn Leu Ala Leu Ala Asp Leu Leu Ser Val Ile Trp Phe
115 120 125

Pro Leu Lys Ile Ala Tyr His Ile His Ala Asn Asn Trp Ile Tyr Gly
130 135 140

Glu Ala Leu Cys Asn Val Leu Ile Gly Phe Phe Tyr Gly Asn Met Tyr
145 150 155 160

Cys Ser Ile Leu Phe Met Thr Cys Leu Ser Val Gln Arg Tyr Trp Val
165 170 175

Ile Val Asn Pro Met Gly His Ser Arg Lys Lys Ala Asn Ile Ala Ile
180 185 190

Gly Ile Ser Leu Ala Ile Trp Leu Leu Ile Leu Leu Val Thr Ile Pro
195 200 205

Leu Tyr Val Val Lys Gln Thr Ile Phe Ile Pro Ala Leu Asn Ile Thr
210 215 220

Thr Cys His Asp Val Leu Pro Glu Gln Leu Leu Val Gly Asp Met Phe
225 230 235 240

Asn Tyr Phe Leu Ser Leu Ala Ile Gly Val Phe Leu Phe Pro Ala Phe
245 250 255

Leu Thr Ala Ser Ala Tyr Val Leu Met Ile Arg Met Leu Arg Ser Ser
260 265 270

Ala Met Asp Glu Asn Ser Glu Lys Lys Arg Lys Arg Ala Ile Lys Leu
275 280 285

Ile Val Thr Val Leu Ala Met Tyr Leu Ile Cys Phe Thr Pro Ser Asn
290 295 300

Leu Leu Leu Val Val His Tyr Phe Leu Ile Lys Ser Gln Gly Gln Ser
305 310 315 320

His Val Tyr Ala Leu Tyr Ile Val Ala Leu Cys Leu Ser Thr Leu Asn
325 330 335

Ser Cys Ile Asp Pro Phe Val Tyr Tyr Phe Val Ser His Asp Phe Arg
340 345 350

Asp His Ala Lys Asn Ala Leu Leu Cys Arg Ser Val Arg Thr Val Lys
355 360 365

Gln Met Gln Val Ser Leu Thr Ser Lys Lys His Ser Arg Lys Ser Ser
370 375 380

Ser Tyr Ser Ser Ser Ser Thr Thr Val Lys Thr Ser Tyr
385 390 395

<210> 89
<211> 1560
<212> PRT
<213> Homo Sapiens

<400> 89

Met Pro Cys Ala Gln Arg Ser Trp Leu Ala Asn Leu Ser Val Val Ala
1 5 10 15

Gln Leu Leu Asn Phe Gly Ala Leu Cys Tyr Gly Arg Gln Pro Gln Pro
20 25 30

Gly Pro Val Arg Phe Pro Asp Arg Arg Gln Glu His Phe Ile Lys Gly
35 40 45

Leu Pro Glu Tyr His Val Val Gly Pro Val Arg Val Asp Ala Ser Gly
50 55 60

His Phe Leu Ser Tyr Gly Leu His Tyr Pro Ile Thr Ser Ser Arg Arg
65 70 75 80

Lys Arg Asp Leu Asp Gly Ser Glu Asp Trp Val Tyr Tyr Arg Ile Ser
85 90 95

His Glu Glu Lys Asp Leu Phe Phe Asn Leu Thr Val Asn Gln Gly Phe
100 105 110

Leu Ser Asn Ser Tyr Ile Met Glu Lys Arg Tyr Gly Asn Leu Ser His
115 120 125

Val Lys Met Met Ala Ser Ser Ala Pro Leu Cys His Leu Ser Gly Thr
130 135 140

Val Leu Gln Gln Gly Thr Arg Val Gly Thr Ala Ala Leu Ser Ala Cys
145 150 155 160

His Gly Leu Thr Gly Phe Phe Gln Leu Pro His Gly Asp Phe Phe Ile

165 170 175

Glu Pro Val Lys Lys His Pro Leu Val Glu Gly Gly Tyr His Pro His
180 185 190

Ile Val Tyr Arg Arg Gln Lys Val Pro Glu Thr Lys Glu Pro Thr Cys
195 200 205

Gly Leu Lys Asp Ser Val Asn Ile Ser Gln Lys Gln Glu Leu Trp Arg
210 215 220

Glu Lys Trp Glu Arg His Asn Leu Pro Ser Arg Ser Leu Ser Arg Arg
225 230 235 240

Ser Ile Ser Lys Glu Arg Trp Val Glu Thr Leu Val Val Ala Asp Thr
245 250 255

Lys Met Ile Glu Tyr His Gly Ser Glu Asn Val Glu Ser Tyr Ile Leu
260 265 270

Thr Ile Met Asn Met Val Thr Gly Leu Phe His Asn Pro Ser Ile Gly
275 280 285

Asn Ala Ile His Ile Val Val Val Arg Leu Ile Leu Leu Glu Glu Glu
290 295 300

Glu Gln Gly Leu Lys Ile Val His His Ala Glu Lys Thr Leu Ser Ser
305 310 315 320

Phe Cys Lys Trp Gln Lys Ser Ile Asn Pro Lys Ser Asp Leu Asn Pro
325 330 335

Val His His Asp Val Ala Val Leu Leu Thr Arg Lys Asp Ile Cys Ala
340 345 350

Gly Phe Asn Arg Pro Cys Glu Thr Leu Gly Leu Ser His Leu Ser Gly
355 360 365

Met Cys Gln Pro His Arg Ser Cys Asn Ile Asn Glu Asp Ser Gly Leu
370 375 380

Pro Leu Ala Phe Thr Ile Ala His Glu Leu Gly His Ser Phe Gly Ile
385 390 395 400

Gln His Asp Gly Lys Glu Asn Asp Cys Glu Pro Val Gly Arg His Pro
405 410 415

```
Tyr Ile Met Ser Arg Gln Leu Gln Tyr Asp Pro Thr Pro Leu Thr Trp
            420                 425                 430

Ser Lys Cys Ser Glu Glu Tyr Ile Thr Arg Phe Leu Asp Arg Gly Trp
        435                 440                 445

Gly Phe Cys Leu Asp Asp Ile Pro Lys Lys Lys Gly Leu Lys Ser Lys
    450                 455                 460

Val Ile Ala Pro Gly Val Ile Tyr Asp Val His His Gln Cys Gln Leu
465                 470                 475                 480

Gln Tyr Gly Pro Asn Ala Thr Phe Cys Gln Glu Val Glu Asn Val Cys
                485                 490                 495

Gln Thr Leu Trp Cys Ser Val Lys Gly Phe Cys Arg Ser Lys Leu Asp
            500                 505                 510

Ala Ala Ala Asp Gly Thr Gln Cys Gly Glu Lys Lys Trp Cys Met Ala
        515                 520                 525

Gly Lys Cys Ile Thr Val Gly Lys Lys Pro Glu Ser Ile Pro Gly Gly
    530                 535                 540

Trp Gly Arg Trp Ser Pro Trp Ser His Cys Ser Arg Thr Cys Gly Ala
545                 550                 555                 560

Gly Val Gln Ser Ala Glu Arg Leu Cys Asn Asn Pro Glu Pro Lys Phe
                565                 570                 575

Gly Gly Lys Tyr Cys Thr Gly Glu Arg Lys Arg Tyr Arg Leu Cys Asn
            580                 585                 590

Val His Pro Cys Arg Ser Glu Ala Pro Thr Phe Arg Gln Met Gln Cys
        595                 600                 605

Ser Glu Phe Asp Thr Val Pro Tyr Lys Asn Glu Leu Tyr His Trp Phe
    610                 615                 620

Pro Ile Phe Asn Pro Ala His Pro Cys Glu Leu Tyr Cys Arg Pro Ile
625                 630                 635                 640

Asp Gly Gln Phe Ser Glu Lys Met Leu Asp Ala Val Ile Asp Gly Thr
                645                 650                 655
```

Pro Cys Phe Glu Gly Gly Asn Ser Arg Asn Val Cys Ile Asn Gly Ile
660 665 670

Cys Lys Met Val Gly Cys Asp Tyr Glu Ile Asp Ser Asn Ala Thr Glu
675 680 685

Asp Arg Cys Gly Val Cys Leu Gly Asp Gly Ser Ser Cys Gln Thr Val
690 695 700

Arg Lys Met Phe Lys Gln Lys Glu Gly Ser Gly Tyr Val Asp Ile Gly
705 710 715 720

Leu Ile Pro Lys Gly Ala Arg Asp Ile Arg Val Met Glu Ile Glu Gly
725 730 735

Ala Gly Asn Phe Leu Ala Ile Arg Ser Glu Asp Pro Glu Lys Tyr Tyr
740 745 750

Leu Asn Gly Gly Phe Ile Ile Gln Trp Asn Gly Asn Tyr Lys Leu Ala
755 760 765

Gly Thr Val Phe Gln Tyr Asp Arg Lys Gly Asp Leu Glu Lys Leu Met
770 775 780

Ala Thr Gly Pro Thr Asn Glu Ser Val Trp Ile Gln Leu Leu Phe Gln
785 790 795 800

Val Thr Asn Pro Gly Ile Lys Tyr Glu Tyr Thr Ile Gln Lys Asp Gly
805 810 815

Leu Asp Asn Asp Val Glu Gln Met Tyr Phe Trp Gln Tyr Gly His Trp
820 825 830

Thr Glu Cys Ser Val Thr Cys Gly Thr Gly Ile Arg Arg Gln Thr Ala
835 840 845

His Cys Ile Lys Lys Gly Arg Gly Met Val Lys Ala Thr Phe Cys Asp
850 855 860

Pro Glu Thr Gln Pro Asn Gly Arg Gln Lys Lys Cys His Glu Lys Ala
865 870 875 880

Cys Pro Pro Arg Trp Trp Ala Gly Glu Trp Glu Ala Cys Ser Ala Thr
885 890 895

```
Cys Gly Pro His Gly Glu Lys Lys Arg Thr Val Leu Cys Ile Gln Thr
            900                 905                 910

Met Val Ser Asp Glu Gln Ala Leu Pro Pro Thr Asp Cys Gln His Leu
        915                 920                 925

Leu Lys Pro Lys Thr Leu Leu Ser Cys Asn Arg Asp Ile Leu Cys Pro
    930                 935                 940

Ser Asp Trp Thr Val Gly Asn Trp Ser Glu Cys Ser Val Ser Cys Gly
945                 950                 955                 960

Gly Gly Val Arg Ile Arg Ser Val Thr Cys Ala Lys Asn His Asp Glu
                965                 970                 975

Pro Cys Asp Val Thr Arg Lys Pro Asn Ser Arg Ala Leu Cys Gly Leu
            980                 985                 990

Gln Gln Cys Pro Ser Ser Arg Arg  Val Leu Lys Pro Asn  Lys Gly Thr
        995                 1000                1005

Ile Ser  Asn Gly Lys Asn Pro  Pro Thr Leu Lys Pro  Val Pro Pro
    1010                1015                1020

Pro Thr  Ser Arg Pro Arg Met  Leu Thr Thr Pro Thr  Gly Pro Glu
    1025                1030                1035

Ser Met  Ser Thr Ser Thr Pro  Ala Ile Ser Ser Pro  Ser Pro Thr
    1040                1045                1050

Thr Ala  Ser Lys Glu Gly Asp  Leu Gly Gly Lys Gln  Trp Gln Asp
    1055                1060                1065

Ser Ser  Thr Gln Pro Glu Leu  Ser Ser Arg Tyr Leu  Ile Ser Thr
    1070                1075                1080

Gly Ser  Thr Ser Gln Pro Ile  Leu Thr Ser Gln Ser  Leu Ser Ile
    1085                1090                1095

Gln Pro  Ser Glu Glu Asn Val  Ser Ser Ser Asp Thr  Gly Pro Thr
    1100                1105                1110

Ser Glu  Gly Gly Leu Val Ala  Thr Thr Thr Ser Gly  Ser Gly Leu
    1115                1120                1125

Ser Ser  Ser Arg Asn Pro Ile  Thr Trp Pro Val Thr  Pro Phe Tyr
```

1130 1135 1140

Asn Thr Leu Thr Lys Gly Pro Glu Met Glu Ile His Ser Gly Ser
1145 1150 1155

Gly Glu Glu Arg Glu Gln Pro Glu Asp Lys Asp Glu Ser Asn Pro
1160 1165 1170

Val Ile Trp Thr Lys Ile Arg Val Pro Gly Asn Asp Ala Pro Val
1175 1180 1185

Glu Ser Thr Glu Met Pro Leu Ala Pro Pro Leu Thr Pro Asp Leu
1190 1195 1200

Ser Arg Glu Ser Trp Trp Pro Pro Phe Ser Thr Val Met Glu Gly
1205 1210 1215

Leu Leu Pro Ser Gln Arg Pro Thr Thr Ser Glu Thr Gly Thr Pro
1220 1225 1230

Arg Val Glu Gly Met Val Thr Glu Lys Pro Ala Asn Thr Leu Leu
1235 1240 1245

Pro Leu Gly Gly Asp His Gln Pro Glu Pro Ser Gly Lys Thr Ala
1250 1255 1260

Asn Arg Asn His Leu Lys Leu Pro Asn Asn Met Asn Gln Thr Lys
1265 1270 1275

Ser Ser Glu Pro Val Leu Thr Glu Glu Asp Ala Thr Ser Leu Ile
1280 1285 1290

Thr Glu Gly Phe Leu Leu Asn Ala Ser Asn Tyr Lys Gln Leu Thr
1295 1300 1305

Asn Gly His Gly Ser Ala His Trp Ile Val Gly Asn Trp Ser Glu
1310 1315 1320

Cys Ser Thr Thr Cys Gly Leu Gly Ala Tyr Trp Lys Arg Val Glu
1325 1330 1335

Cys Thr Thr Gln Met Asp Ser Asp Cys Ala Ala Ile Gln Arg Pro
1340 1345 1350

Asp Pro Ala Lys Arg Cys His Leu Arg Pro Cys Ala Gly Trp Lys
1355 1360 1365

Val Gly Asn Trp Ser Lys Cys Ser Arg Asn Cys Ser Gly Gly Phe
1370 1375 1380

Lys Ile Arg Glu Ile Gln Cys Val Asp Ser Arg Asp His Arg Asn
1385 1390 1395

Leu Arg Pro Phe His Cys Gln Phe Leu Ala Gly Ile Pro Pro Pro
1400 1405 1410

Leu Ser Met Ser Cys Asn Pro Glu Pro Cys Glu Ala Trp Gln Val
1415 1420 1425

Glu Pro Trp Ser Gln Cys Ser Arg Ser Cys Gly Gly Gly Val Gln
1430 1435 1440

Glu Arg Gly Val Phe Cys Pro Gly Gly Leu Cys Asp Trp Thr Lys
1445 1450 1455

Arg Pro Thr Ser Thr Met Ser Cys Asn Glu His Leu Cys Cys His
1460 1465 1470

Trp Ala Thr Gly Asn Trp Asp Leu Cys Ser Thr Ser Cys Gly Gly
1475 1480 1485

Gly Phe Gln Lys Arg Ile Val Gln Cys Val Pro Ser Glu Gly Asn
1490 1495 1500

Lys Thr Glu Asp Gln Asp Gln Cys Leu Cys Asp His Lys Pro Arg
1505 1510 1515

Pro Pro Glu Phe Lys Lys Cys Asn Gln Gln Ala Cys Lys Lys Ser
1520 1525 1530

Ala Asp Leu Leu Cys Thr Lys Asp Lys Leu Ser Ala Ser Phe Cys
1535 1540 1545

Gln Thr Leu Lys Ala Met Lys Lys Cys Ser Val Pro
1550 1555 1560

<210> 90
<211> 96
<212> PRT
<213> Homo Sapiens

<400> 90

Met Cys Cys Thr Lys Ser Leu Leu Leu Ala Ala Leu Met Ser Val Leu
1 5 10 15

Leu Leu His Leu Cys Gly Glu Ser Glu Ala Ala Ser Asn Phe Asp Cys
20 25 30

Cys Leu Gly Tyr Thr Asp Arg Ile Leu His Pro Lys Phe Ile Val Gly
35 40 45

Phe Thr Arg Gln Leu Ala Asn Glu Gly Cys Asp Ile Asn Ala Ile Ile
50 55 60

Phe His Thr Lys Lys Lys Leu Ser Val Cys Ala Asn Pro Lys Gln Thr
65 70 75 80

Trp Val Lys Tyr Ile Val Arg Leu Leu Ser Lys Lys Val Lys Asn Met
85 90 95

<210> 91
<211> 336
<212> PRT
<213> Homo Sapiens

<400> 91

Met Leu Gln Ser Leu Ala Gly Ser Ser Cys Val Arg Leu Val Glu Arg
1 5 10 15

His Arg Ser Ala Trp Cys Phe Gly Phe Leu Val Leu Gly Tyr Leu Leu
20 25 30

Tyr Leu Val Phe Gly Ala Val Val Phe Ser Ser Val Glu Leu Pro Tyr
35 40 45

Glu Asp Leu Leu Arg Gln Glu Leu Arg Lys Leu Lys Arg Arg Phe Leu
50 55 60

Glu Glu His Glu Cys Leu Ser Glu Gln Gln Leu Glu Gln Phe Leu Gly
65 70 75 80

Arg Val Leu Glu Ala Ser Asn Tyr Gly Val Ser Val Leu Ser Asn Ala
85 90 95

Ser Gly Asn Trp Asn Trp Asp Phe Thr Ser Ala Leu Phe Phe Ala Ser
100 105 110

Thr Val Leu Ser Thr Thr Gly Tyr Gly His Thr Val Pro Leu Ser Asp
115 120 125

```
Gly Gly Lys Ala Phe Cys Ile Ile Tyr Ser Val Ile Gly Ile Pro Phe
    130                 135                 140

Thr Leu Leu Phe Leu Thr Ala Val Val Gln Arg Ile Thr Val His Val
145                 150                 155                 160

Thr Arg Arg Pro Val Leu Tyr Phe His Ile Arg Trp Gly Phe Ser Lys
                165                 170                 175

Gln Val Val Ala Ile Val His Ala Val Leu Leu Gly Phe Val Thr Val
            180                 185                 190

Ser Cys Phe Phe Phe Ile Pro Ala Ala Val Phe Ser Val Leu Glu Asp
        195                 200                 205

Asp Trp Asn Phe Leu Glu Ser Phe Tyr Phe Cys Phe Ile Ser Leu Ser
    210                 215                 220

Thr Ile Gly Leu Gly Asp Tyr Val Pro Gly Glu Gly Tyr Asn Gln Lys
225                 230                 235                 240

Phe Arg Glu Leu Tyr Lys Ile Gly Ile Thr Cys Tyr Leu Leu Leu Gly
                245                 250                 255

Leu Ile Ala Met Leu Val Val Leu Glu Thr Phe Cys Glu Leu His Glu
            260                 265                 270

Leu Lys Lys Phe Arg Lys Met Phe Tyr Val Lys Lys Asp Lys Asp Glu
        275                 280                 285

Asp Gln Val His Ile Ile Glu His Asp Gln Leu Ser Phe Ser Ser Ile
    290                 295                 300

Thr Asp Gln Ala Ala Gly Met Lys Glu Asp Gln Lys Gln Asn Glu Pro
305                 310                 315                 320

Phe Val Ala Thr Gln Ser Ser Ala Cys Val Asp Gly Pro Ala Asn His
                325                 330                 335

<210>  92
<211>  103
<212>  PRT
<213>  Homo Sapiens

<400>  92
```

Met Glu Thr Thr Asn Gly Thr Glu Thr Trp Tyr Glu Ser Leu His Ala
1 5 10 15

Val Leu Lys Ala Leu Asn Ala Thr Leu His Ser Asn Leu Leu Cys Arg
20 25 30

Pro Gly Pro Gly Leu Gly Pro Asp Asn Gln Thr Glu Glu Arg Arg Ala
35 40 45

Ser Leu Pro Gly Arg Asp Asp Asn Ser Tyr Met Tyr Ile Leu Phe Val
50 55 60

Met Phe Leu Phe Ala Val Thr Val Gly Ser Leu Ile Leu Gly Tyr Thr
65 70 75 80

Arg Ser Arg Lys Val Asp Lys Arg Ser Asp Pro Tyr His Val Tyr Ile
85 90 95

Lys Asn Arg Val Ser Met Ile
100

<210> 93
<211> 4590
<212> PRT
<213> Homo Sapiens

<400> 93

Met Gly Arg His Leu Ala Leu Leu Leu Leu Leu Leu Leu Leu Phe Gln
1 5 10 15

His Phe Gly Asp Ser Asp Gly Ser Gln Arg Leu Glu Gln Thr Pro Leu
20 25 30

Gln Phe Thr His Leu Glu Tyr Asn Val Thr Val Gln Glu Asn Ser Ala
35 40 45

Ala Lys Thr Tyr Val Gly His Pro Val Lys Met Gly Val Tyr Ile Thr
50 55 60

His Pro Ala Trp Glu Val Arg Tyr Lys Ile Val Ser Gly Asp Ser Glu
65 70 75 80

Asn Leu Phe Lys Ala Glu Glu Tyr Ile Leu Gly Asp Phe Cys Phe Leu
85 90 95

Arg Ile Arg Thr Lys Gly Gly Asn Thr Ala Ile Leu Asn Arg Glu Val
100 105 110

```
Lys Asp His Tyr Thr Leu Ile Val Lys Ala Leu Glu Lys Asn Thr Asn
        115                 120                 125

Val Glu Ala Arg Thr Lys Val Arg Val Gln Val Leu Asp Thr Asn Asp
    130                 135                 140

Leu Arg Pro Leu Phe Ser Pro Thr Ser Tyr Ser Val Ser Leu Pro Glu
145                 150                 155                 160

Asn Thr Ala Ile Arg Thr Ser Ile Ala Arg Val Ser Ala Thr Asp Ala
                165                 170                 175

Asp Ile Gly Thr Asn Gly Glu Phe Tyr Tyr Ser Phe Lys Asp Arg Thr
            180                 185                 190

Asp Met Phe Ala Ile His Pro Thr Ser Gly Val Ile Val Leu Thr Gly
        195                 200                 205

Arg Leu Asp Tyr Leu Glu Thr Lys Leu Tyr Glu Met Glu Ile Leu Ala
    210                 215                 220

Ala Asp Arg Gly Met Lys Leu Tyr Gly Ser Ser Gly Ile Ser Ser Met
225                 230                 235                 240

Ala Lys Leu Thr Val His Ile Glu Gln Ala Asn Glu Cys Ala Pro Val
                245                 250                 255

Ile Thr Ala Val Thr Leu Ser Pro Ser Glu Leu Asp Arg Asp Pro Ala
            260                 265                 270

Tyr Ala Ile Val Thr Val Asp Asp Cys Asp Gln Gly Ala Asn Gly Asp
        275                 280                 285

Ile Ala Ser Leu Ser Ile Val Ala Gly Asp Leu Leu Gln Gln Phe Arg
    290                 295                 300

Thr Val Arg Ser Phe Pro Gly Ser Lys Glu Tyr Lys Val Lys Ala Ile
305                 310                 315                 320

Gly Asp Ile Asp Trp Asp Ser His Pro Phe Gly Tyr Asn Leu Thr Leu
                325                 330                 335

Gln Ala Lys Asp Lys Gly Thr Pro Pro Gln Phe Ser Ser Val Lys Val
            340                 345                 350
```

Ile His Val Thr Ser Pro Gln Phe Lys Ala Gly Pro Val Lys Phe Glu
355 360 365

Lys Asp Val Tyr Arg Ala Glu Ile Ser Glu Phe Ala Pro Pro Asn Thr
370 375 380

Pro Val Val Met Val Lys Ala Ile Pro Ala Tyr Ser His Leu Arg Tyr
385 390 395 400

Val Phe Lys Arg Thr Pro Gly Lys Ala Lys Phe Ser Leu Asn Tyr Asn
405 410 415

Thr Gly Leu Ile Ser Ile Leu Glu Pro Val Lys Arg Gln Gln Ala Ala
420 425 430

His Phe Glu Leu Glu Val Thr Thr Ser Asp Arg Lys Ala Ser Thr Lys
435 440 445

Val Leu Val Lys Val Leu Gly Ala Asn Ser Asn Pro Pro Glu Phe Thr
450 455 460

Gln Thr Ala Tyr Lys Ala Ala Phe Asp Glu Asn Val Pro Ile Gly Thr
465 470 475 480

Thr Ile Met Ser Leu Ser Ala Val Asp Pro Asp Glu Gly Glu Asn Gly
485 490 495

Tyr Val Thr Tyr Ser Ile Ala Asn Leu Asn His Val Pro Phe Ala Ile
500 505 510

Asp His Phe Thr Gly Ala Val Ser Thr Ser Glu Asn Leu Asp Tyr Glu
515 520 525

Leu Met Pro Arg Val Tyr Thr Leu Arg Ile Arg Ala Ser Asp Trp Gly
530 535 540

Leu Pro Tyr Arg Arg Glu Val Glu Val Leu Ala Thr Ile Thr Leu Asn
545 550 555 560

Asn Leu Asn Asp Asn Thr Pro Leu Phe Glu Lys Ile Asn Cys Glu Gly
565 570 575

Thr Ile Pro Arg Asp Leu Gly Val Gly Glu Gln Ile Thr Thr Val Ser
580 585 590

Ala Ile Asp Ala Asp Glu Leu Gln Leu Val Gln Tyr Gln Ile Glu Ala
595 600 605

Gly Asn Glu Leu Asp Leu Phe Ser Leu Asn Pro Asn Ser Gly Val Leu
610 615 620

Ser Leu Lys Arg Ser Leu Met Asp Gly Leu Gly Ala Lys Val Ser Phe
625 630 635 640

His Ser Leu Arg Ile Thr Ala Thr Asp Gly Glu Asn Phe Ala Thr Pro
645 650 655

Leu Tyr Ile Asn Ile Thr Val Ala Ala Ser His Lys Leu Val Asn Leu
660 665 670

Gln Cys Glu Glu Thr Gly Val Ala Lys Met Leu Ala Glu Lys Leu Leu
675 680 685

Gln Ala Asn Lys Leu His Asn Gln Gly Glu Val Glu Asp Ile Phe Phe
690 695 700

Asp Ser His Ser Val Asn Ala His Ile Pro Gln Phe Arg Ser Thr Leu
705 710 715 720

Pro Thr Gly Ile Gln Val Lys Glu Asn Gln Pro Val Gly Ser Ser Val
725 730 735

Ile Phe Met Asn Ser Thr Asp Leu Asp Thr Gly Phe Asn Gly Lys Leu
740 745 750

Val Tyr Ala Val Ser Gly Gly Asn Glu Asp Ser Cys Phe Met Ile Asp
755 760 765

Met Glu Thr Gly Met Leu Lys Ile Leu Ser Pro Leu Asp Arg Glu Thr
770 775 780

Thr Asp Lys Tyr Thr Leu Asn Ile Thr Val Tyr Asp Leu Gly Ile Pro
785 790 795 800

Gln Lys Ala Ala Trp Arg Leu Leu His Val Val Val Val Asp Ala Asn
805 810 815

Asp Asn Pro Pro Glu Phe Leu Gln Glu Ser Tyr Phe Val Glu Val Ser
820 825 830

Glu Asp Lys Glu Val His Ser Glu Ile Ile Gln Val Glu Ala Thr Asp

835 840 845

Lys Asp Leu Gly Pro Asn Gly His Val Thr Tyr Ser Ile Leu Thr Asp
850 855 860

Thr Asp Thr Phe Ser Ile Asp Ser Val Thr Gly Val Val Asn Ile Ala
865 870 875 880

Arg Pro Leu Asp Arg Glu Leu Gln His Glu His Ser Leu Lys Ile Glu
885 890 895

Ala Arg Asp Gln Ala Arg Glu Glu Pro Gln Leu Phe Ser Thr Val Val
900 905 910

Val Lys Val Ser Leu Glu Asp Val Asn Asp Asn Pro Pro Thr Phe Ile
915 920 925

Pro Pro Asn Tyr Arg Val Lys Val Arg Glu Asp Leu Pro Glu Gly Thr
930 935 940

Val Ile Met Trp Leu Glu Ala His Asp Pro Asp Leu Gly Gln Ser Gly
945 950 955 960

Gln Val Arg Tyr Ser Leu Leu Asp His Gly Glu Gly Asn Phe Asp Val
965 970 975

Asp Lys Leu Ser Gly Ala Val Arg Ile Val Gln Gln Leu Asp Phe Glu
980 985 990

Lys Lys Gln Val Tyr Asn Leu Thr Val Arg Ala Lys Asp Lys Gly Lys
995 1000 1005

Pro Val Ser Leu Ser Ser Thr Cys Tyr Val Glu Val Glu Val Val
1010 1015 1020

Asp Val Asn Glu Asn Leu His Pro Pro Val Phe Ser Ser Phe Val
1025 1030 1035

Glu Lys Gly Thr Val Lys Glu Asp Ala Pro Val Gly Ser Leu Val
1040 1045 1050

Met Thr Val Ser Ala His Asp Glu Asp Ala Gly Arg Asp Gly Glu
1055 1060 1065

Ile Arg Tyr Ser Ile Arg Asp Gly Ser Gly Val Gly Val Phe Lys
1070 1075 1080

Ile Gly Glu Glu Thr Gly Val Ile Glu Thr Ser Asp Arg Leu Asp
1085 1090 1095

Arg Glu Ser Thr Ser His Tyr Trp Leu Thr Val Phe Ala Thr Asp
1100 1105 1110

Gln Gly Val Val Pro Leu Ser Ser Phe Ile Glu Ile Tyr Ile Glu
1115 1120 1125

Val Glu Asp Val Asn Asp Asn Ala Pro Gln Thr Ser Glu Pro Val
1130 1135 1140

Tyr Tyr Pro Glu Ile Met Glu Asn Ser Pro Lys Asp Val Ser Val
1145 1150 1155

Val Gln Ile Glu Ala Phe Asp Pro Asp Ser Ser Ser Asn Asp Lys
1160 1165 1170

Leu Met Tyr Lys Ile Thr Ser Gly Asn Pro Gln Gly Phe Phe Ser
1175 1180 1185

Ile His Pro Lys Thr Gly Leu Ile Thr Thr Thr Ser Arg Lys Leu
1190 1195 1200

Asp Arg Glu Gln Gln Asp Glu His Ile Leu Glu Val Thr Val Thr
1205 1210 1215

Asp Asn Gly Ser Pro Pro Lys Ser Thr Ile Ala Arg Val Ile Val
1220 1225 1230

Lys Ile Leu Asp Glu Asn Asp Asn Lys Pro Gln Phe Leu Gln Lys
1235 1240 1245

Phe Tyr Lys Ile Arg Leu Pro Glu Arg Glu Lys Pro Asp Arg Glu
1250 1255 1260

Arg Asn Ala Arg Arg Glu Pro Leu Tyr Arg Val Ile Ala Thr Asp
1265 1270 1275

Lys Asp Glu Gly Pro Asn Ala Glu Ile Ser Tyr Ser Ile Glu Asp
1280 1285 1290

Gly Asn Glu His Gly Lys Phe Phe Ile Glu Pro Lys Thr Gly Val
1295 1300 1305

Val Ser Ser Lys Arg Phe Ser Ala Ala Gly Glu Tyr Asp Ile Leu
1310 1315 1320

Ser Ile Lys Ala Val Asp Asn Gly Arg Pro Gln Lys Ser Ser Thr
1325 1330 1335

Thr Arg Leu His Ile Glu Trp Ile Ser Lys Pro Lys Gln Ser Leu
1340 1345 1350

Glu Pro Ile Ser Phe Glu Glu Ser Phe Phe Thr Phe Thr Val Met
1355 1360 1365

Glu Ser Asp Pro Val Ala His Met Ile Gly Val Ile Ser Val Glu
1370 1375 1380

Pro Pro Gly Ile Pro Leu Trp Phe Asp Ile Thr Gly Gly Asn Tyr
1385 1390 1395

Asp Ser His Phe Asp Val Asp Lys Gly Thr Gly Thr Ile Ile Val
1400 1405 1410

Ala Lys Pro Leu Asp Ala Glu Gln Lys Ser Asn Tyr Asn Leu Thr
1415 1420 1425

Val Glu Ala Thr Asp Gly Thr Thr Thr Ile Leu Thr Gln Val Phe
1430 1435 1440

Ile Lys Val Ile Asp Thr Asn Asp His Arg Pro Gln Phe Ser Thr
1445 1450 1455

Ser Lys Tyr Glu Val Val Ile Pro Glu Asp Thr Ala Pro Glu Thr
1460 1465 1470

Glu Ile Leu Gln Ile Ser Ala Val Asp Gln Asp Glu Lys Asn Lys
1475 1480 1485

Leu Ile Tyr Thr Leu Gln Ser Ser Arg Asp Pro Leu Ser Leu Lys
1490 1495 1500

Lys Phe Arg Leu Asp Pro Ala Thr Gly Ser Leu Tyr Thr Ser Glu
1505 1510 1515

Lys Leu Asp His Glu Ala Val Ser Pro Ala His Leu Thr Val Met
1520 1525 1530

Val Arg Asp Gln Asp Val Pro Val Lys Arg Asn Phe Ala Arg Ile
1535 1540 1545

Val Val Asn Val Ser Asp Thr Asn Asp His Ala Pro Trp Phe Thr
1550 1555 1560

Ala Ser Ser Tyr Lys Gly Arg Val Tyr Glu Ser Ala Ala Val Gly
1565 1570 1575

Ser Val Val Leu Gln Val Thr Ala Leu Asp Lys Asp Lys Gly Lys
1580 1585 1590

Asn Ala Glu Val Leu Tyr Ser Ile Glu Ser Gly Asn Ile Gly Asn
1595 1600 1605

Ile Gly Asn Ser Phe Met Ile Asp Pro Val Leu Gly Ser Ile Lys
1610 1615 1620

Thr Ala Lys Glu Leu Asp Arg Ser Asn Gln Ala Glu Tyr Asp Leu
1625 1630 1635

Met Val Lys Ala Thr Asp Lys Gly Ser Pro Pro Met Ser Glu Ile
1640 1645 1650

Thr Ser Val Arg Ile Phe Val Thr Ile Ala Asp Asn Ala Ser Pro
1655 1660 1665

Lys Phe Thr Ser Lys Glu Tyr Ser Val Glu Leu Ser Glu Thr Val
1670 1675 1680

Ser Ile Gly Ser Phe Val Gly Met Val Thr Ala His Ser Gln Ser
1685 1690 1695

Ser Val Val Tyr Glu Ile Lys Asp Gly Asn Thr Gly Asp Ala Phe
1700 1705 1710

Asp Ile Asn Pro His Ser Gly Thr Ile Ile Thr Gln Lys Ala Leu
1715 1720 1725

Asp Phe Glu Thr Leu Pro Ile Tyr Thr Leu Ile Ile Gln Gly Thr
1730 1735 1740

Asn Met Ala Gly Leu Ser Thr Asn Thr Thr Val Leu Val His Leu
1745 1750 1755

Gln Asp Glu Asn Asp Asn Ala Pro Val Phe Met Gln Ala Glu Tyr

1760 1765 1770

Thr Gly Leu Ile Ser Glu Ser Ala Ser Ile Asn Ser Val Val Leu
1775 1780 1785

Thr Asp Arg Asn Val Pro Leu Val Ile Arg Ala Ala Asp Ala Asp
1790 1795 1800

Lys Asp Ser Asn Ala Leu Leu Val Tyr His Ile Val Glu Pro Ser
1805 1810 1815

Val His Thr Tyr Phe Ala Ile Asp Ser Ser Thr Gly Ala Ile His
1820 1825 1830

Thr Val Leu Ser Leu Asp Tyr Glu Glu Thr Ser Ile Phe His Phe
1835 1840 1845

Thr Val Gln Val His Asp Met Gly Thr Pro Arg Leu Phe Ala Glu
1850 1855 1860

Tyr Ala Ala Asn Val Thr Val His Val Ile Asp Ile Asn Asp Cys
1865 1870 1875

Pro Pro Val Phe Ala Lys Pro Leu Tyr Glu Ala Ser Leu Leu Leu
1880 1885 1890

Pro Thr Tyr Lys Gly Val Lys Val Ile Thr Val Asn Ala Thr Asp
1895 1900 1905

Ala Asp Ser Ser Ala Phe Ser Gln Leu Ile Tyr Ser Ile Thr Glu
1910 1915 1920

Gly Asn Ile Gly Glu Lys Phe Ser Met Asp Tyr Lys Thr Gly Ala
1925 1930 1935

Leu Thr Val Gln Asn Thr Thr Gln Leu Arg Ser Arg Tyr Glu Leu
1940 1945 1950

Thr Val Arg Ala Ser Asp Gly Arg Phe Ala Gly Leu Thr Ser Val
1955 1960 1965

Lys Ile Asn Val Lys Glu Ser Lys Glu Ser His Leu Lys Phe Thr
1970 1975 1980

Gln Asp Val Tyr Ser Ala Val Val Lys Glu Asn Ser Thr Glu Ala
1985 1990 1995

Glu Thr Leu Ala Val Ile Thr Ala Ile Gly Ser Pro Ile Asn Glu
2000 2005 2010

Pro Leu Phe Tyr His Ile Leu Asn Pro Asp Arg Arg Phe Lys Ile
2015 2020 2025

Ser Arg Thr Ser Gly Val Leu Ser Thr Thr Gly Thr Pro Phe Asp
2030 2035 2040

Arg Glu Gln Gln Glu Ala Phe Asp Val Val Val Glu Val Ile Glu
2045 2050 2055

Glu His Lys Pro Ser Ala Val Ala His Val Val Val Lys Val Ile
2060 2065 2070

Val Glu Asp Gln Asn Asp Asn Ala Pro Val Phe Val Asn Leu Pro
2075 2080 2085

Tyr Tyr Ala Val Val Lys Val Asp Thr Glu Val Gly His Val Ile
2090 2095 2100

Arg Tyr Val Thr Ala Val Asp Arg Asp Ser Gly Arg Asn Gly Glu
2105 2110 2115

Val His Tyr Tyr Leu Lys Glu His His Glu His Phe Gln Ile Gly
2120 2125 2130

Pro Leu Gly Glu Ile Ser Leu Lys Lys Gln Phe Glu Leu Asp Thr
2135 2140 2145

Leu Asn Lys Glu Tyr Leu Val Thr Val Val Ala Lys Asp Gly Gly
2150 2155 2160

Asn Pro Ala Phe Ser Ala Glu Val Ile Val Pro Ile Thr Val Met
2165 2170 2175

Asn Lys Ala Met Pro Val Phe Glu Lys Pro Phe Tyr Ser Ala Glu
2180 2185 2190

Ile Ala Glu Ser Ile Gln Val His Ser Pro Val Val His Val Gln
2195 2200 2205

Ala Asn Ser Pro Glu Gly Leu Lys Val Phe Tyr Ser Ile Thr Asp
2210 2215 2220

Gly Asp Pro Phe Ser Gln Phe Thr Ile Asn Phe Asn Thr Gly Val
2225 2230 2235

Ile Asn Val Ile Ala Pro Leu Asp Phe Glu Ala His Pro Ala Tyr
2240 2245 2250

Lys Leu Ser Ile Arg Ala Thr Asp Ser Leu Thr Gly Ala His Ala
2255 2260 2265

Glu Val Phe Val Asp Ile Ile Val Asp Asp Ile Asn Asp Asn Pro
2270 2275 2280

Pro Val Phe Ala Gln Gln Ser Tyr Ala Val Thr Leu Ser Glu Ala
2285 2290 2295

Ser Val Ile Gly Thr Ser Val Val Gln Val Arg Ala Thr Asp Ser
2300 2305 2310

Asp Ser Glu Pro Asn Arg Gly Ile Ser Tyr Gln Met Phe Gly Asn
2315 2320 2325

His Ser Lys Ser His Asp His Phe His Val Asp Ser Ser Thr Gly
2330 2335 2340

Leu Ile Ser Leu Leu Arg Thr Leu Asp Tyr Glu Gln Ser Arg Gln
2345 2350 2355

His Thr Ile Phe Val Arg Ala Val Asp Gly Gly Met Pro Thr Leu
2360 2365 2370

Ser Ser Asp Val Ile Val Thr Val Asp Val Thr Asp Leu Asn Gly
2375 2380 2385

Asn Pro Pro Leu Phe Glu Gln Gln Ile Tyr Glu Ala Arg Ile Ser
2390 2395 2400

Glu His Ala Pro His Gly His Phe Val Thr Cys Val Lys Ala Tyr
2405 2410 2415

Asp Ala Asp Ser Ser Asp Ile Asp Lys Leu Gln Tyr Ser Ile Leu
2420 2425 2430

Ser Gly Asn Asp His Lys His Phe Val Ile Asp Ser Ala Thr Gly
2435 2440 2445

Ile Ile Thr Leu Ser Asn Leu His Arg His Ala Leu Lys Pro Phe
2450 2455 2460

Tyr Ser Leu Asn Leu Ser Val Ser Asp Gly Val Phe Arg Ser Ser
2465 2470 2475

Thr Gln Val His Val Thr Val Ile Gly Gly Asn Leu His Ser Pro
2480 2485 2490

Ala Phe Leu Gln Asn Glu Tyr Glu Val Glu Leu Ala Glu Asn Ala
2495 2500 2505

Pro Leu His Thr Leu Val Met Glu Val Lys Thr Thr Asp Gly Asp
2510 2515 2520

Ser Gly Ile Tyr Gly His Val Thr Tyr His Ile Val Asn Asp Phe
2525 2530 2535

Ala Lys Asp Arg Phe Tyr Ile Asn Glu Arg Gly Gln Ile Phe Thr
2540 2545 2550

Leu Glu Lys Leu Asp Arg Glu Thr Pro Ala Glu Lys Val Ile Ser
2555 2560 2565

Val Arg Leu Met Ala Lys Asp Ala Gly Gly Lys Val Ala Phe Cys
2570 2575 2580

Thr Val Asn Val Ile Leu Thr Asp Asp Asn Asp Asn Ala Pro Gln
2585 2590 2595

Phe Arg Ala Thr Lys Tyr Glu Val Asn Ile Gly Ser Ser Ala Ala
2600 2605 2610

Lys Gly Thr Ser Val Val Lys Ser Ala Ser Asp Ala Asp Glu Gly
2615 2620 2625

Ser Asn Ala Asp Ile Thr Tyr Ala Ile Glu Ala Asp Ser Glu Ser
2630 2635 2640

Val Lys Glu Asn Leu Glu Ile Asn Lys Leu Ser Gly Val Ile Thr
2645 2650 2655

Thr Lys Glu Ser Leu Ile Gly Leu Glu Asn Glu Phe Phe Thr Phe
2660 2665 2670

Phe Val Arg Ala Val Asp Asn Gly Ser Pro Ser Lys Glu Ser Val

2675 2680 2685

Val Leu Val Tyr Val Lys Ile Leu Pro Pro Glu Met Gln Leu Pro
2690 2695 2700

Lys Phe Ser Glu Pro Phe Tyr Thr Phe Thr Val Ser Glu Asp Val
2705 2710 2715

Pro Val Gly Thr Glu Ile Asp Leu Ile Arg Ala Glu His Ser Gly
2720 2725 2730

Thr Val Leu Tyr Ser Leu Val Lys Gly Asn Thr Pro Glu Ser Asn
2735 2740 2745

Arg Asp Glu Ser Phe Val Ile Asp Arg Gln Ser Gly Arg Leu Lys
2750 2755 2760

Leu Glu Lys Ser Leu Asp His Glu Thr Thr Lys Trp Tyr Gln Phe
2765 2770 2775

Ser Ile Leu Ala Arg Cys Thr Gln Asp Asp His Glu Met Val Ala
2780 2785 2790

Ser Val Asp Val Ser Ile Gln Val Lys Asp Ala Asn Asp Asn Ser
2795 2800 2805

Pro Val Phe Glu Ser Ser Pro Tyr Glu Ala Phe Ile Val Glu Asn
2810 2815 2820

Leu Pro Gly Gly Ser Arg Val Ile Gln Ile Arg Ala Ser Asp Ala
2825 2830 2835

Asp Ser Gly Thr Asn Gly Gln Val Met Tyr Ser Leu Asp Gln Ser
2840 2845 2850

Gln Ser Val Glu Val Ile Glu Ser Phe Ala Ile Asn Met Glu Thr
2855 2860 2865

Gly Trp Ile Thr Thr Leu Lys Glu Leu Asp His Glu Lys Arg Asp
2870 2875 2880

Asn Tyr Gln Ile Lys Val Val Ala Ser Asp His Gly Glu Lys Ile
2885 2890 2895

Gln Leu Ser Ser Thr Ala Ile Val Asp Val Thr Val Thr Asp Val
2900 2905 2910

Asn Asp Ser Pro Pro Arg Phe Thr Ala Glu Ile Tyr Lys Gly Thr
2915 2920 2925

Val Ser Glu Asp Asp Pro Gln Gly Gly Val Ile Ala Ile Leu Ser
2930 2935 2940

Thr Thr Asp Ala Asp Ser Glu Glu Ile Asn Arg Gln Val Thr Tyr
2945 2950 2955

Phe Ile Thr Gly Gly Asp Pro Leu Gly Gln Phe Ala Val Glu Thr
2960 2965 2970

Ile Gln Asn Glu Trp Lys Val Tyr Val Lys Lys Pro Leu Asp Arg
2975 2980 2985

Glu Lys Arg Asp Asn Tyr Leu Leu Thr Ile Thr Ala Thr Asp Gly
2990 2995 3000

Thr Phe Ser Ser Lys Ala Ile Val Glu Val Lys Val Leu Asp Ala
3005 3010 3015

Asn Asp Asn Ser Pro Val Cys Glu Lys Thr Leu Tyr Ser Asp Thr
3020 3025 3030

Ile Pro Glu Asp Val Leu Pro Gly Lys Leu Ile Met Gln Ile Ser
3035 3040 3045

Ala Thr Asp Ala Asp Ile Arg Ser Asn Ala Glu Ile Thr Tyr Thr
3050 3055 3060

Leu Leu Gly Ser Gly Ala Glu Lys Phe Lys Leu Asn Pro Asp Thr
3065 3070 3075

Gly Glu Leu Lys Thr Ser Thr Pro Leu Asp Arg Glu Glu Gln Ala
3080 3085 3090

Val Tyr His Leu Leu Val Arg Ala Thr Asp Gly Gly Gly Arg Phe
3095 3100 3105

Cys Gln Ala Ser Ile Val Val Thr Leu Glu Asp Val Asn Asp Asn
3110 3115 3120

Ala Pro Glu Phe Ser Ala Asp Pro Tyr Ala Ile Thr Val Phe Glu
3125 3130 3135

Asn Thr Glu Pro Gly Thr Leu Leu Thr Arg Val Gln Ala Thr Asp
3140 3145 3150

Ala Asp Ala Gly Leu Asn Arg Lys Ile Leu Tyr Ser Leu Ile Asp
3155 3160 3165

Ser Ala Asp Gly Gln Phe Ser Ile Asn Glu Leu Ser Gly Ile Ile
3170 3175 3180

Gln Leu Glu Lys Pro Leu Asp Arg Glu Leu Gln Ala Val Tyr Thr
3185 3190 3195

Leu Ser Leu Lys Ala Val Asp Gln Gly Leu Pro Arg Arg Leu Thr
3200 3205 3210

Ala Thr Gly Thr Val Ile Val Ser Val Leu Asp Ile Asn Asp Asn
3215 3220 3225

Pro Pro Val Phe Glu Tyr Arg Glu Tyr Gly Ala Thr Val Ser Glu
3230 3235 3240

Asp Ile Leu Val Gly Thr Glu Val Leu Gln Val Tyr Ala Ala Ser
3245 3250 3255

Arg Asp Ile Glu Ala Asn Ala Glu Ile Thr Tyr Ser Ile Ile Ser
3260 3265 3270

Gly Asn Glu His Gly Lys Phe Ser Ile Asp Ser Lys Thr Gly Ala
3275 3280 3285

Val Phe Ile Ile Glu Asn Leu Asp Tyr Glu Ser Ser His Glu Tyr
3290 3295 3300

Tyr Leu Thr Val Glu Ala Thr Asp Gly Gly Thr Pro Ser Leu Ser
3305 3310 3315

Asp Val Ala Thr Val Asn Val Asn Val Thr Asp Ile Asn Asp Asn
3320 3325 3330

Thr Pro Val Phe Ser Gln Asp Thr Tyr Thr Thr Val Ile Ser Glu
3335 3340 3345

Asp Ala Val Leu Glu Gln Ser Val Ile Thr Val Met Ala Asp Asp
3350 3355 3360

Ala Asp Gly Pro Ser Asn Ser His Ile His Tyr Ser Ile Ile Asp
3365 3370 3375

Gly Asn Gln Gly Ser Ser Phe Thr Ile Asp Pro Val Arg Gly Glu
3380 3385 3390

Val Lys Val Thr Lys Leu Leu Asp Arg Glu Thr Ile Ser Gly Tyr
3395 3400 3405

Thr Leu Thr Val Gln Ala Ser Asp Asn Gly Ser Pro Pro Arg Val
3410 3415 3420

Asn Thr Thr Thr Val Asn Ile Asp Val Ser Asp Val Asn Asp Asn
3425 3430 3435

Ala Pro Val Phe Ser Arg Gly Asn Tyr Ser Val Ile Ile Gln Glu
3440 3445 3450

Asn Lys Pro Val Gly Phe Ser Val Leu Gln Leu Val Val Thr Asp
3455 3460 3465

Glu Asp Ser Ser His Asn Gly Pro Pro Phe Phe Phe Thr Ile Val
3470 3475 3480

Thr Gly Asn Asp Glu Lys Ala Phe Glu Val Asn Pro Gln Gly Val
3485 3490 3495

Leu Leu Thr Ser Ser Ala Ile Lys Arg Lys Glu Lys Asp His Tyr
3500 3505 3510

Leu Leu Gln Val Lys Val Ala Asp Asn Gly Lys Pro Gln Leu Ser
3515 3520 3525

Ser Leu Thr Tyr Ile Asp Ile Arg Val Ile Glu Glu Ser Ile Tyr
3530 3535 3540

Pro Pro Ala Ile Leu Pro Leu Glu Ile Phe Ile Thr Ser Ser Gly
3545 3550 3555

Glu Glu Tyr Ser Gly Gly Val Ile Gly Lys Ile His Ala Thr Asp
3560 3565 3570

Gln Asp Val Tyr Asp Thr Leu Thr Tyr Ser Leu Asp Pro Gln Met
3575 3580 3585

Asp Asn Leu Phe Ser Val Ser Ser Thr Gly Gly Lys Leu Ile Ala

3590 3595 3600

His Lys Lys Leu Asp Ile Gly Gln Tyr Leu Leu Asn Val Ser Val
3605 3610 3615

Thr Asp Gly Lys Phe Thr Thr Val Ala Asp Ile Thr Val His Ile
3620 3625 3630

Arg Gln Val Thr Gln Glu Met Leu Asn His Thr Ile Ala Ile Arg
3635 3640 3645

Phe Ala Asn Leu Thr Pro Glu Glu Phe Val Gly Asp Tyr Trp Arg
3650 3655 3660

Asn Phe Gln Arg Ala Leu Arg Asn Ile Leu Gly Val Arg Arg Asn
3665 3670 3675

Asp Ile Gln Ile Val Ser Leu Gln Ser Ser Glu Pro His Pro His
3680 3685 3690

Leu Asp Val Leu Leu Phe Val Glu Lys Pro Gly Ser Ala Gln Ile
3695 3700 3705

Ser Thr Lys Gln Leu Leu His Lys Ile Asn Ser Ser Val Thr Asp
3710 3715 3720

Ile Glu Glu Ile Ile Gly Val Arg Ile Leu Asn Val Phe Gln Lys
3725 3730 3735

Leu Cys Ala Gly Leu Asp Cys Pro Trp Lys Phe Cys Asp Glu Lys
3740 3745 3750

Val Ser Val Asp Glu Ser Val Met Ser Thr His Ser Thr Ala Arg
3755 3760 3765

Leu Ser Phe Val Thr Pro Arg His His Arg Ala Ala Val Cys Leu
3770 3775 3780

Cys Lys Glu Gly Arg Cys Pro Pro Val His His Gly Cys Glu Asp
3785 3790 3795

Asp Pro Cys Pro Glu Gly Ser Glu Cys Val Ser Asp Pro Trp Glu
3800 3805 3810

Glu Lys His Thr Cys Val Cys Pro Ser Gly Arg Phe Gly Gln Cys
3815 3820 3825

Pro Gly Ser Ser Ser Met Thr Leu Thr Gly Asn Ser Tyr Val Lys
3830 3835 3840

Tyr Arg Leu Thr Glu Asn Glu Asn Lys Leu Glu Met Lys Leu Thr
3845 3850 3855

Met Arg Leu Arg Thr Tyr Ser Thr His Ala Val Val Met Tyr Ala
3860 3865 3870

Arg Gly Thr Asp Tyr Ser Ile Leu Glu Ile His His Gly Arg Leu
3875 3880 3885

Gln Tyr Lys Phe Asp Cys Gly Ser Gly Pro Gly Ile Val Ser Val
3890 3895 3900

Gln Ser Ile Gln Val Asn Asp Gly Gln Trp His Ala Val Ala Leu
3905 3910 3915

Glu Val Asn Gly Asn Tyr Ala Arg Leu Val Leu Asp Gln Val His
3920 3925 3930

Thr Ala Ser Gly Thr Ala Pro Gly Thr Leu Lys Thr Leu Asn Leu
3935 3940 3945

Asp Asn Tyr Val Phe Phe Gly Gly His Ile Arg Gln Gln Gly Thr
3950 3955 3960

Arg His Gly Arg Ser Pro Gln Val Gly Asn Gly Phe Arg Gly Cys
3965 3970 3975

Met Asp Ser Ile Tyr Leu Asn Gly Gln Glu Leu Pro Leu Asn Ser
3980 3985 3990

Lys Pro Arg Ser Tyr Ala His Ile Glu Glu Ser Val Asp Val Ser
3995 4000 4005

Pro Gly Cys Phe Leu Thr Ala Thr Glu Asp Cys Ala Ser Asn Pro
4010 4015 4020

Cys Gln Asn Gly Gly Val Cys Asn Pro Ser Pro Ala Gly Gly Tyr
4025 4030 4035

Tyr Cys Lys Cys Ser Ala Leu Tyr Ile Gly Thr His Cys Glu Ile
4040 4045 4050

Ser Val Asn Pro Cys Ser Ser Asn Pro Cys Leu Tyr Gly Gly Thr
4055 4060 4065

Cys Val Val Asp Asn Gly Gly Phe Val Cys Gln Cys Arg Gly Leu
4070 4075 4080

Tyr Thr Gly Gln Arg Cys Gln Leu Ser Pro Tyr Cys Lys Asp Glu
4085 4090 4095

Pro Cys Lys Asn Gly Gly Thr Cys Phe Asp Ser Leu Asp Gly Ala
4100 4105 4110

Val Cys Gln Cys Asp Ser Gly Phe Arg Gly Glu Arg Cys Gln Ser
4115 4120 4125

Asp Ile Asp Glu Cys Ser Gly Asn Pro Cys Leu His Gly Ala Leu
4130 4135 4140

Cys Glu Asn Thr His Gly Ser Tyr His Cys Asn Cys Ser His Glu
4145 4150 4155

Tyr Arg Gly Arg His Cys Glu Asp Ala Ala Pro Asn Gln Tyr Val
4160 4165 4170

Ser Thr Pro Trp Asn Ile Gly Leu Ala Glu Gly Ile Gly Ile Val
4175 4180 4185

Val Phe Val Ala Gly Ile Phe Leu Leu Val Val Val Phe Val Leu
4190 4195 4200

Cys Arg Lys Met Ile Ser Arg Lys Lys Lys His Gln Ala Glu Pro
4205 4210 4215

Lys Asp Lys His Leu Gly Pro Ala Thr Ala Phe Leu Gln Arg Pro
4220 4225 4230

Tyr Phe Asp Ser Lys Leu Asn Lys Asn Ile Tyr Ser Asp Ile Pro
4235 4240 4245

Pro Gln Val Pro Val Arg Pro Ile Ser Tyr Thr Pro Ser Ile Pro
4250 4255 4260

Ser Asp Ser Arg Asn Asn Leu Asp Arg Asn Ser Phe Glu Gly Ser
4265 4270 4275

Ala Ile Pro Glu His Pro Glu Phe Ser Thr Phe Asn Pro Glu Ser
4280 4285 4290

Val His Gly His Arg Lys Ala Val Ala Val Cys Ser Val Ala Pro
4295 4300 4305

Asn Leu Pro Pro Pro Pro Pro Ser Asn Ser Pro Ser Asp Ser Asp
4310 4315 4320

Ser Ile Gln Lys Pro Ser Trp Asp Phe Asp Tyr Asp Thr Lys Val
4325 4330 4335

Val Asp Leu Asp Pro Cys Leu Ser Lys Lys Pro Leu Glu Glu Lys
4340 4345 4350

Pro Ser Gln Pro Tyr Ser Ala Arg Glu Ser Leu Ser Glu Val Gln
4355 4360 4365

Ser Leu Ser Ser Phe Gln Ser Glu Ser Cys Asp Asp Asn Gly Tyr
4370 4375 4380

His Trp Asp Thr Ser Asp Trp Met Pro Ser Val Pro Leu Pro Asp
4385 4390 4395

Ile Gln Glu Phe Pro Asn Tyr Glu Val Ile Asp Glu Gln Thr Pro
4400 4405 4410

Leu Tyr Ser Ala Asp Pro Asn Ala Ile Asp Thr Asp Tyr Tyr Pro
4415 4420 4425

Gly Gly Tyr Asp Ile Glu Ser Asp Phe Pro Pro Pro Pro Glu Asp
4430 4435 4440

Phe Pro Ala Ala Asp Glu Leu Pro Pro Leu Pro Pro Glu Phe Ser
4445 4450 4455

Asn Gln Phe Glu Ser Ile His Pro Pro Arg Asp Met Pro Ala Ala
4460 4465 4470

Gly Ser Leu Gly Ser Ser Ser Arg Asn Arg Gln Arg Phe Asn Leu
4475 4480 4485

Asn Gln Tyr Leu Pro Asn Phe Tyr Pro Leu Asp Met Ser Glu Pro
4490 4495 4500

Gln Thr Lys Gly Thr Gly Glu Asn Ser Thr Cys Arg Glu Pro His

4505 4510 4515

Ala Pro Tyr Pro Pro Gly Tyr Gln Arg His Phe Glu Ala Pro Ala
4520 4525 4530

Val Glu Ser Met Pro Met Ser Val Tyr Ala Ser Thr Ala Ser Cys
4535 4540 4545

Ser Asp Val Ser Ala Cys Cys Glu Val Glu Ser Glu Val Met Met
4550 4555 4560

Ser Asp Tyr Glu Ser Gly Asp Asp Gly His Phe Glu Glu Val Thr
4565 4570 4575

Ile Pro Pro Leu Asp Ser Gln Gln His Thr Glu Val
4580 4585 4590

<210> 94
<211> 202
<212> PRT
<213> Homo Sapiens

<400> 94

Met Cys Tyr Gly Lys Cys Ala Arg Cys Ile Gly His Ser Leu Val Gly
1 5 10 15

Leu Ala Leu Leu Cys Ile Ala Ala Asn Ile Leu Leu Tyr Phe Pro Asn
20 25 30

Gly Glu Thr Lys Tyr Ala Ser Glu Asn His Leu Ser Arg Phe Val Trp
35 40 45

Phe Phe Ser Gly Ile Val Gly Gly Gly Leu Leu Met Leu Leu Pro Ala
50 55 60

Phe Val Phe Ile Gly Leu Glu Gln Asp Asp Cys Cys Gly Cys Cys Gly
65 70 75 80

His Glu Asn Cys Gly Lys Arg Cys Ala Met Leu Ser Ser Val Leu Ala
85 90 95

Ala Leu Ile Gly Ile Ala Gly Ser Gly Tyr Cys Val Ile Val Ala Ala
100 105 110

Leu Gly Leu Ala Glu Gly Pro Leu Cys Leu Asp Ser Leu Gly Gln Trp
115 120 125

Asn Tyr Thr Phe Ala Ser Thr Glu Gly Gln Tyr Leu Leu Asp Thr Ser
130 135 140

Thr Trp Ser Glu Cys Thr Glu Pro Lys His Ile Val Glu Trp Asn Val
145 150 155 160

Ser Leu Phe Ser Ile Leu Leu Ala Leu Gly Gly Ile Glu Phe Ile Leu
165 170 175

Cys Leu Ile Gln Val Ile Asn Gly Val Leu Gly Gly Ile Cys Gly Phe
180 185 190

Cys Cys Ser His Gln Gln Gln Tyr Asp Cys
195 200

<210> 95
<211> 1035
<212> PRT
<213> Homo Sapiens

<400> 95

Met Ser Thr Glu Asn Val Glu Gly Lys Pro Ser Asn Leu Gly Glu Arg
1 5 10 15

Gly Arg Ala Arg Ser Ser Thr Phe Leu Arg Val Val Gln Pro Met Phe
20 25 30

Asn His Ser Ile Phe Thr Ser Ala Val Ser Pro Ala Ala Glu Arg Ile
35 40 45

Arg Phe Ile Leu Gly Glu Glu Asp Asp Ser Pro Ala Pro Pro Gln Leu
50 55 60

Phe Thr Glu Leu Asp Glu Leu Leu Ala Val Asp Gly Gln Glu Met Glu
65 70 75 80

Trp Lys Glu Thr Ala Arg Trp Ile Lys Phe Glu Glu Lys Val Glu Gln
85 90 95

Gly Gly Glu Arg Trp Ser Lys Pro His Val Ala Thr Leu Ser Leu His
100 105 110

Ser Leu Phe Glu Leu Arg Thr Cys Met Glu Lys Gly Ser Ile Met Leu
115 120 125

Asp Arg Glu Ala Ser Ser Leu Pro Gln Leu Val Glu Met Ile Val Asp

130 135 140

His Gln Ile Glu Thr Gly Leu Leu Lys Pro Glu Leu Lys Asp Lys Val
145 150 155 160

Thr Tyr Thr Leu Leu Arg Lys His Arg His Gln Thr Lys Lys Ser Asn
165 170 175

Leu Arg Ser Leu Ala Asp Ile Gly Lys Thr Val Ser Ser Ala Ser Arg
180 185 190

Met Phe Thr Asn Pro Asp Asn Gly Ser Pro Ala Met Thr His Arg Asn
195 200 205

Leu Thr Ser Ser Ser Leu Asn Asp Ile Ser Asp Lys Pro Glu Lys Asp
210 215 220

Gln Leu Lys Asn Lys Phe Met Lys Lys Leu Pro Arg Asp Ala Glu Ala
225 230 235 240

Ser Asn Val Leu Val Gly Glu Val Asp Phe Leu Asp Thr Pro Phe Ile
245 250 255

Ala Phe Val Arg Leu Gln Gln Ala Val Met Leu Gly Ala Leu Thr Glu
260 265 270

Val Pro Val Pro Thr Arg Phe Leu Phe Ile Leu Leu Gly Pro Lys Gly
275 280 285

Lys Ala Lys Ser Tyr His Glu Ile Gly Arg Ala Ile Ala Thr Leu Met
290 295 300

Ser Asp Glu Val Phe His Asp Ile Ala Tyr Lys Ala Lys Asp Arg His
305 310 315 320

Asp Leu Ile Ala Gly Ile Asp Glu Phe Leu Asp Glu Val Ile Val Leu
325 330 335

Pro Pro Gly Glu Trp Asp Pro Ala Ile Arg Ile Glu Pro Pro Lys Ser
340 345 350

Leu Pro Ser Ser Asp Lys Arg Lys Asn Met Tyr Ser Gly Gly Glu Asn
355 360 365

Val Gln Met Asn Gly Asp Thr Pro His Asp Gly Gly His Gly Gly Gly
370 375 380

Gly His Gly Asp Cys Glu Glu Leu Gln Arg Thr Gly Arg Phe Cys Gly
385 390 395 400

Gly Leu Ile Lys Asp Ile Lys Arg Lys Ala Pro Phe Phe Ala Ser Asp
405 410 415

Phe Tyr Asp Ala Leu Asn Ile Gln Ala Leu Ser Ala Ile Leu Phe Ile
420 425 430

Tyr Leu Ala Thr Val Thr Asn Ala Ile Thr Phe Gly Gly Leu Leu Gly
435 440 445

Asp Ala Thr Asp Asn Met Gln Gly Val Leu Glu Ser Phe Leu Gly Thr
450 455 460

Ala Val Ser Gly Ala Ile Phe Cys Leu Phe Ala Gly Gln Pro Leu Thr
465 470 475 480

Ile Leu Ser Ser Thr Gly Pro Val Leu Val Phe Glu Arg Leu Leu Phe
485 490 495

Asn Phe Ser Lys Asp Asn Asn Phe Asp Tyr Leu Glu Phe Arg Leu Trp
500 505 510

Ile Gly Leu Trp Ser Ala Phe Leu Cys Leu Ile Leu Val Ala Thr Asp
515 520 525

Ala Ser Phe Leu Val Gln Tyr Phe Thr Arg Phe Thr Glu Glu Gly Phe
530 535 540

Ser Ser Leu Ile Ser Phe Ile Phe Ile Tyr Asp Ala Phe Lys Lys Met
545 550 555 560

Ile Lys Leu Ala Asp Tyr Tyr Pro Ile Asn Ser Asn Phe Lys Val Gly
565 570 575

Tyr Asn Thr Leu Phe Ser Cys Thr Cys Val Pro Pro Asp Pro Ala Asn
580 585 590

Ile Ser Ile Ser Asn Asp Thr Thr Leu Ala Pro Glu Tyr Leu Pro Thr
595 600 605

Met Ser Ser Thr Asp Met Tyr His Asn Thr Thr Phe Asp Trp Ala Phe
610 615 620

Leu Ser Lys Lys Glu Cys Ser Lys Tyr Gly Gly Asn Leu Val Gly Asn
625 630 635 640

Asn Cys Asn Phe Val Pro Asp Ile Thr Leu Met Ser Phe Ile Leu Phe
645 650 655

Leu Gly Thr Tyr Thr Ser Ser Met Ala Leu Lys Lys Phe Lys Thr Ser
660 665 670

Pro Tyr Phe Pro Thr Thr Ala Arg Lys Leu Ile Ser Asp Phe Ala Ile
675 680 685

Ile Leu Ser Ile Leu Ile Phe Cys Val Ile Asp Ala Leu Val Gly Val
690 695 700

Asp Thr Pro Lys Leu Ile Val Pro Ser Glu Phe Lys Pro Thr Ser Pro
705 710 715 720

Asn Arg Gly Trp Phe Val Pro Pro Phe Gly Glu Asn Pro Trp Trp Val
725 730 735

Cys Leu Ala Ala Ala Ile Pro Ala Leu Leu Val Thr Ile Leu Ile Phe
740 745 750

Met Asp Gln Gln Ile Thr Ala Val Ile Val Asn Arg Lys Glu His Lys
755 760 765

Leu Lys Lys Gly Ala Gly Tyr His Leu Asp Leu Phe Trp Val Ala Ile
770 775 780

Leu Met Val Ile Cys Ser Leu Met Ala Leu Pro Trp Tyr Val Ala Ala
785 790 795 800

Thr Val Ile Ser Ile Ala His Ile Asp Ser Leu Lys Met Glu Thr Glu
805 810 815

Thr Ser Ala Pro Gly Glu Gln Pro Lys Phe Leu Gly Val Arg Glu Gln
820 825 830

Arg Val Thr Gly Thr Leu Val Phe Ile Leu Thr Gly Leu Ser Val Phe
835 840 845

Met Ala Pro Ile Leu Lys Phe Ile Pro Met Pro Val Leu Tyr Gly Val
850 855 860

Phe Leu Tyr Met Gly Val Ala Ser Leu Asn Gly Val Gln Phe Met Asp
865 870 875 880

Arg Leu Lys Leu Leu Leu Met Pro Leu Lys His Gln Pro Asp Phe Ile
885 890 895

Tyr Leu Arg His Val Pro Leu Arg Arg Val His Leu Phe Thr Phe Leu
900 905 910

Gln Val Leu Cys Leu Ala Leu Leu Trp Ile Leu Lys Ser Thr Val Ala
915 920 925

Ala Ile Ile Phe Pro Val Met Ile Leu Ala Leu Val Ala Val Arg Lys
930 935 940

Gly Met Asp Tyr Leu Phe Ser Gln His Asp Leu Ser Phe Leu Asp Asp
945 950 955 960

Val Ile Pro Glu Lys Asp Lys Lys Lys Lys Glu Asp Glu Lys Lys Lys
965 970 975

Lys Lys Lys Lys Gly Ser Leu Asp Ser Asp Asn Asp Asp Ser Asp Cys
980 985 990

Pro Tyr Ser Glu Lys Val Pro Ser Ile Lys Ile Pro Met Asp Ile Met
995 1000 1005

Glu Gln Gln Pro Phe Leu Ser Asp Ser Lys Pro Ser Asp Arg Glu
1010 1015 1020

Arg Ser Pro Thr Phe Leu Glu Arg His Thr Ser Cys
1025 1030 1035

<210> 96
<211> 480
<212> PRT
<213> Homo Sapiens

<400> 96

Met Ser Thr Pro Gly Val Asn Ser Ser Ala Ser Leu Ser Pro Asp Arg
1 5 10 15

Leu Asn Ser Pro Val Thr Ile Pro Ala Val Met Phe Ile Phe Gly Val
20 25 30

Val Gly Asn Leu Val Ala Ile Val Val Leu Cys Lys Ser Arg Lys Glu
35 40 45

Gln Lys Glu Thr Thr Phe Tyr Thr Leu Val Cys Gly Leu Ala Val Thr
50 55 60

Asp Leu Leu Gly Thr Leu Leu Val Ser Pro Val Thr Ile Ala Thr Tyr
65 70 75 80

Met Lys Gly Gln Trp Pro Gly Gly Gln Pro Leu Cys Glu Tyr Ser Thr
85 90 95

Phe Ile Leu Leu Phe Phe Ser Leu Ser Gly Leu Ser Ile Ile Cys Ala
100 105 110

Met Ser Val Glu Arg Tyr Leu Ala Ile Asn His Ala Tyr Phe Tyr Ser
115 120 125

His Tyr Val Asp Lys Arg Leu Ala Gly Leu Thr Leu Phe Ala Val Tyr
130 135 140

Ala Ser Asn Val Leu Phe Cys Ala Leu Pro Asn Met Gly Leu Gly Ser
145 150 155 160

Ser Arg Leu Gln Tyr Pro Asp Thr Trp Cys Phe Ile Asp Trp Thr Thr
165 170 175

Asn Val Thr Ala His Ala Ala Tyr Ser Tyr Met Tyr Ala Gly Phe Ser
180 185 190

Ser Phe Leu Ile Leu Ala Thr Val Leu Cys Asn Val Leu Val Cys Gly
195 200 205

Ala Leu Leu Arg Met His Arg Gln Phe Met Arg Arg Thr Ser Leu Gly
210 215 220

Thr Glu Gln His His Ala Ala Ala Ala Ala Ser Val Ala Ser Arg Gly
225 230 235 240

His Pro Ala Ala Ser Pro Ala Leu Pro Arg Leu Ser Asp Phe Arg Arg
245 250 255

Arg Arg Ser Phe Arg Arg Ile Ala Gly Ala Glu Ile Gln Met Val Ile
260 265 270

Leu Leu Ile Ala Thr Ser Leu Val Val Leu Ile Cys Ser Ile Pro Leu
275 280 285

Val Val Arg Val Phe Val Asn Gln Leu Tyr Gln Pro Ser Leu Glu Arg
290 295 300

Glu Val Ser Lys Asn Pro Asp Leu Gln Ala Ile Arg Ile Ala Ser Val
305 310 315 320

Asn Pro Ile Leu Asp Pro Trp Ile Tyr Ile Leu Leu Arg Lys Thr Val
325 330 335

Leu Ser Lys Ala Ile Glu Lys Ile Lys Cys Leu Phe Cys Arg Ile Gly
340 345 350

Gly Ser Arg Arg Glu Arg Ser Gly Gln His Cys Ser Asp Ser Gln Arg
355 360 365

Thr Ser Ser Ala Met Ser Gly His Ser Arg Ser Phe Ile Ser Arg Glu
370 375 380

Leu Lys Glu Ile Ser Ser Thr Ser Gln Thr Leu Leu Pro Asp Leu Ser
385 390 395 400

Leu Pro Asp Leu Ser Glu Asn Gly Leu Gly Gly Arg Asn Leu Leu Pro
405 410 415

Gly Val Pro Gly Met Gly Leu Ala Gln Glu Asp Thr Thr Ser Leu Arg
420 425 430

Thr Leu Arg Ile Ser Glu Thr Ser Asp Ser Ser Gln Gly Gln Asp Ser
435 440 445

Glu Ser Val Leu Leu Val Asp Glu Ala Gly Gly Ser Gly Arg Ala Gly
450 455 460

Pro Ala Pro Lys Gly Ser Ser Leu Gln Val Thr Phe Pro Ser Glu Thr
465 470 475 480

<210> 97
<211> 335
<212> PRT
<213> Homo Sapiens

<400> 97

Met Gly His Pro Pro Leu Leu Pro Leu Leu Leu Leu Leu His Thr Cys
1 5 10 15

Val Pro Ala Ser Trp Gly Leu Arg Cys Met Gln Cys Lys Thr Asn Gly

20 25 30

Asp Cys Arg Val Glu Glu Cys Ala Leu Gly Gln Asp Leu Cys Arg Thr
35 40 45

Thr Ile Val Arg Leu Trp Glu Glu Gly Glu Glu Leu Glu Leu Val Glu
50 55 60

Lys Ser Cys Thr His Ser Glu Lys Thr Asn Arg Thr Leu Ser Tyr Arg
65 70 75 80

Thr Gly Leu Lys Ile Thr Ser Leu Thr Glu Val Val Cys Gly Leu Asp
85 90 95

Leu Cys Asn Gln Gly Asn Ser Gly Arg Ala Val Thr Tyr Ser Arg Ser
100 105 110

Arg Tyr Leu Glu Cys Ile Ser Cys Gly Ser Ser Asp Met Ser Cys Glu
115 120 125

Arg Gly Arg His Gln Ser Leu Gln Cys Arg Ser Pro Glu Glu Gln Cys
130 135 140

Leu Asp Val Val Thr His Trp Ile Gln Glu Gly Glu Glu Gly Arg Pro
145 150 155 160

Lys Asp Asp Arg His Leu Arg Gly Cys Gly Tyr Leu Pro Gly Cys Pro
165 170 175

Gly Ser Asn Gly Phe His Asn Asn Asp Thr Phe His Phe Leu Lys Cys
180 185 190

Cys Asn Thr Thr Lys Cys Asn Glu Gly Pro Ile Leu Glu Leu Glu Asn
195 200 205

Leu Pro Gln Asn Gly Arg Gln Cys Tyr Ser Cys Lys Gly Asn Ser Thr
210 215 220

His Gly Cys Ser Ser Glu Glu Thr Phe Leu Ile Asp Cys Arg Gly Pro
225 230 235 240

Met Asn Gln Cys Leu Val Ala Thr Gly Thr His Glu Pro Lys Asn Gln
245 250 255

Ser Tyr Met Val Arg Gly Cys Ala Thr Ala Ser Met Cys Gln His Ala
260 265 270

His Leu Gly Asp Ala Phe Ser Met Asn His Ile Asp Val Ser Cys Cys
275 280 285

Thr Lys Ser Gly Cys Asn His Pro Asp Leu Asp Val Gln Tyr Arg Ser
290 295 300

Gly Ala Ala Pro Gln Pro Gly Pro Ala His Leu Ser Leu Thr Ile Thr
305 310 315 320

Leu Leu Met Thr Ala Arg Leu Trp Gly Gly Thr Leu Leu Trp Thr
325 330 335

<210> 98
<211> 512
<212> PRT
<213> *Homo Sapiens*

<400> 98

Met Asp Phe Glu Ser Gly Gln Val Asp Pro Leu Ala Ser Val Ile Leu
1 5 10 15

Pro Pro Asn Leu Leu Glu Asn Leu Ser Pro Glu Asp Ser Val Leu Val
20 25 30

Arg Arg Ala Gln Phe Thr Phe Phe Asn Lys Thr Gly Leu Phe Gln Asp
35 40 45

Val Gly Pro Gln Arg Lys Thr Leu Val Ser Tyr Val Met Ala Cys Ser
50 55 60

Ile Gly Asn Ile Thr Ile Gln Asn Leu Lys Asp Pro Val Gln Ile Lys
65 70 75 80

Ile Lys His Thr Arg Thr Gln Glu Val His His Pro Ile Cys Ala Phe
85 90 95

Trp Asp Leu Asn Lys Asn Lys Ser Phe Gly Gly Trp Asn Thr Ser Gly
100 105 110

Cys Val Ala His Arg Asp Ser Asp Ala Ser Glu Thr Val Cys Leu Cys
115 120 125

Asn His Phe Thr His Phe Gly Val Leu Met Asp Leu Pro Arg Ser Ala
130 135 140

```
Ser Gln Leu Asp Ala Arg Asn Thr Lys Val Leu Thr Phe Ile Ser Tyr
145                 150                 155                 160

Ile Gly Cys Gly Ile Ser Ala Ile Phe Ser Ala Ala Thr Leu Leu Thr
                165                 170                 175

Tyr Val Ala Phe Glu Lys Leu Arg Arg Asp Tyr Pro Ser Lys Ile Leu
            180                 185                 190

Met Asn Leu Ser Thr Ala Leu Leu Phe Leu Asn Leu Leu Phe Leu Leu
        195                 200                 205

Asp Gly Trp Ile Thr Ser Phe Asn Val Asp Gly Leu Cys Ile Ala Val
    210                 215                 220

Ala Val Leu Leu His Phe Phe Leu Leu Ala Thr Phe Thr Trp Met Gly
225                 230                 235                 240

Leu Glu Ala Ile His Met Tyr Ile Ala Leu Val Lys Val Phe Asn Thr
                245                 250                 255

Tyr Ile Arg Arg Tyr Ile Leu Lys Phe Cys Ile Ile Gly Trp Gly Leu
            260                 265                 270

Pro Ala Leu Val Val Ser Val Val Leu Ala Ser Arg Asn Asn Asn Glu
        275                 280                 285

Val Tyr Gly Lys Glu Ser Tyr Gly Lys Glu Lys Gly Asp Glu Phe Cys
    290                 295                 300

Trp Ile Gln Asp Pro Val Ile Phe Tyr Val Thr Cys Ala Gly Tyr Phe
305                 310                 315                 320

Gly Val Met Phe Phe Leu Asn Ile Ala Met Phe Ile Val Val Met Val
                325                 330                 335

Gln Ile Cys Gly Arg Asn Gly Lys Arg Ser Asn Arg Thr Leu Arg Glu
            340                 345                 350

Glu Val Leu Arg Asn Leu Arg Ser Val Val Ser Leu Thr Phe Leu Leu
        355                 360                 365

Gly Met Thr Trp Gly Phe Ala Phe Phe Ala Trp Gly Pro Leu Asn Ile
    370                 375                 380

Pro Phe Met Tyr Leu Phe Ser Ile Phe Asn Ser Leu Gln Gly Leu Phe
```

385 390 395 400

Ile Phe Ile Phe His Cys Ala Met Lys Glu Asn Val Gln Lys Gln Trp
405 410 415

Arg Arg His Leu Cys Cys Gly Arg Phe Arg Leu Ala Asp Asn Ser Asp
420 425 430

Trp Ser Lys Thr Ala Thr Asn Ile Ile Lys Lys Ser Ser Asp Asn Leu
435 440 445

Gly Lys Ser Leu Ser Ser Ser Ser Ile Gly Ser Asn Ser Thr Tyr Leu
450 455 460

Thr Ser Lys Ser Lys Ser Ser Ser Thr Thr Tyr Phe Lys Arg Asn Ser
465 470 475 480

His Thr Asp Asn Val Ser Tyr Glu His Ser Phe Asn Lys Ser Gly Ser
485 490 495

Leu Arg Gln Cys Phe His Gly Gln Val Leu Val Lys Thr Gly Pro Cys
500 505 510

<210> 99
<211> 202
<212> PRT
<213> Homo Sapiens

<400> 99

Met Lys Val Leu Ala Ala Gly Val Val Pro Leu Leu Leu Val Leu His
1 5 10 15

Trp Lys His Gly Ala Gly Ser Pro Leu Pro Ile Thr Pro Val Asn Ala
20 25 30

Thr Cys Ala Ile Arg His Pro Cys His Asn Asn Leu Met Asn Gln Ile
35 40 45

Arg Ser Gln Leu Ala Gln Leu Asn Gly Ser Ala Asn Ala Leu Phe Ile
50 55 60

Leu Tyr Tyr Thr Ala Gln Gly Glu Pro Phe Pro Asn Asn Leu Asp Lys
65 70 75 80

Leu Cys Gly Pro Asn Val Thr Asp Phe Pro Pro Phe His Ala Asn Gly
85 90 95

Thr Glu Lys Ala Lys Leu Val Glu Leu Tyr Arg Ile Val Val Tyr Leu
100 105 110

Gly Thr Ser Leu Gly Asn Ile Thr Arg Asp Gln Lys Ile Leu Asn Pro
115 120 125

Ser Ala Leu Ser Leu His Ser Lys Leu Asn Ala Thr Ala Asp Ile Leu
130 135 140

Arg Gly Leu Leu Ser Asn Val Leu Cys Arg Leu Cys Ser Lys Tyr His
145 150 155 160

Val Gly His Val Asp Val Thr Tyr Gly Pro Asp Thr Ser Gly Lys Asp
165 170 175

Val Phe Gln Lys Lys Lys Leu Gly Cys Gln Leu Leu Gly Lys Tyr Lys
180 185 190

Gln Ile Ile Ala Val Leu Ala Gln Ala Phe
195 200

<210> 100
<211> 504
<212> PRT
<213> Homo Sapiens

<400> 100

Met Thr Pro Ser Pro Leu Leu Leu Leu Leu Leu Pro Pro Leu Leu Leu
1 5 10 15

Gly Ala Phe Pro Pro Ala Ala Ala Ala Arg Gly Pro Pro Lys Met Ala
20 25 30

Asp Lys Val Val Pro Arg Gln Val Ala Arg Leu Gly Arg Thr Val Arg
35 40 45

Leu Gln Cys Pro Val Glu Gly Asp Pro Pro Pro Leu Thr Met Trp Thr
50 55 60

Lys Asp Gly Arg Thr Ile His Ser Gly Trp Ser Arg Phe Arg Val Leu
65 70 75 80

Pro Gln Gly Leu Lys Val Lys Gln Val Glu Arg Glu Asp Ala Gly Val
85 90 95

Tyr Val Cys Lys Ala Thr Asn Gly Phe Gly Ser Leu Ser Val Asn Tyr

100 105 110

Thr Leu Val Val Leu Asp Asp Ile Ser Pro Gly Lys Glu Ser Leu Gly
115 120 125

Pro Asp Ser Ser Ser Gly Gly Gln Glu Asp Pro Ala Ser Gln Gln Trp
130 135 140

Ala Arg Pro Arg Phe Thr Gln Pro Ser Lys Met Arg Arg Arg Val Ile
145 150 155 160

Ala Arg Pro Val Gly Ser Ser Val Arg Leu Lys Cys Val Ala Ser Gly
165 170 175

His Pro Arg Pro Asp Ile Thr Trp Met Lys Asp Asp Gln Ala Leu Thr
180 185 190

Arg Pro Glu Ala Ala Glu Pro Arg Lys Lys Lys Trp Thr Leu Ser Leu
195 200 205

Lys Asn Leu Arg Pro Glu Asp Ser Gly Lys Tyr Thr Cys Arg Val Ser
210 215 220

Asn Arg Ala Gly Ala Ile Asn Ala Thr Tyr Lys Val Asp Val Ile Gln
225 230 235 240

Arg Thr Arg Ser Lys Pro Val Leu Thr Gly Thr His Pro Val Asn Thr
245 250 255

Thr Val Asp Phe Gly Gly Thr Thr Ser Phe Gln Cys Lys Val Arg Ser
260 265 270

Asp Val Lys Pro Val Ile Gln Trp Leu Lys Arg Val Glu Tyr Gly Ala
275 280 285

Glu Gly Arg His Asn Ser Thr Ile Asp Val Gly Gly Gln Lys Phe Val
290 295 300

Val Leu Pro Thr Gly Asp Val Trp Ser Arg Pro Asp Gly Ser Tyr Leu
305 310 315 320

Asn Lys Leu Leu Ile Thr Arg Ala Arg Gln Asp Asp Ala Gly Met Tyr
325 330 335

Ile Cys Leu Gly Ala Asn Thr Met Gly Tyr Ser Phe Arg Ser Ala Phe
340 345 350

Leu Thr Val Leu Pro Asp Pro Lys Pro Gln Gly Pro Pro Val Ala Ser
355 360 365

Ser Ser Ser Ala Thr Ser Leu Pro Trp Pro Val Val Ile Gly Ile Pro
370 375 380

Ala Gly Ala Val Phe Ile Leu Gly Thr Leu Leu Leu Trp Leu Cys Gln
385 390 395 400

Ala Gln Lys Lys Pro Cys Thr Pro Ala Pro Ala Pro Pro Leu Pro Gly
405 410 415

His Arg Pro Pro Gly Thr Ala Arg Asp Arg Ser Gly Asp Lys Asp Leu
420 425 430

Pro Ser Leu Ala Ala Leu Ser Ala Gly Pro Gly Val Gly Leu Cys Glu
435 440 445

Glu His Gly Ser Pro Ala Ala Pro Gln His Leu Leu Gly Pro Gly Pro
450 455 460

Val Ala Gly Pro Lys Leu Tyr Pro Lys Leu Tyr Thr Asp Ile His Thr
465 470 475 480

His Thr His Thr His Ser His Thr His Ser His Val Glu Gly Lys Val
485 490 495

His Gln His Ile His Tyr Gln Cys
500

<210> 101
<211> 915
<212> PRT
<213> Homo Sapiens

<400> 101

Met Gly Arg Pro Arg Leu Thr Leu Val Cys His Val Ser Ile Ile Ile
1 5 10 15

Ser Ala Arg Asp Leu Ser Met Asn Asn Leu Thr Glu Leu Gln Pro Gly
20 25 30

Leu Phe His His Leu Arg Phe Leu Glu Glu Leu Arg Leu Ser Gly Asn
35 40 45

His Leu Ser His Ile Pro Gly Gln Ala Phe Ser Gly Leu Tyr Ser Leu
50 55 60

Lys Ile Leu Met Leu Gln Asn Asn Gln Leu Gly Gly Ile Pro Ala Glu
65 70 75 80

Ala Leu Trp Glu Leu Pro Ser Leu Gln Ser Leu Arg Leu Asp Ala Asn
85 90 95

Leu Ile Ser Leu Val Pro Glu Arg Ser Phe Glu Gly Leu Ser Ser Leu
100 105 110

Arg His Leu Trp Leu Asp Asp Asn Ala Leu Thr Glu Ile Pro Val Arg
115 120 125

Ala Leu Asn Asn Leu Pro Ala Leu Gln Ala Met Thr Leu Ala Leu Asn
130 135 140

Arg Ile Ser His Ile Pro Asp Tyr Ala Phe Gln Asn Leu Thr Ser Leu
145 150 155 160

Val Val Leu His Leu His Asn Asn Arg Ile Gln His Leu Gly Thr His
165 170 175

Ser Phe Glu Gly Leu His Asn Leu Glu Thr Leu Asp Leu Asn Tyr Asn
180 185 190

Lys Leu Gln Glu Phe Pro Val Ala Ile Arg Thr Leu Gly Arg Leu Gln
195 200 205

Glu Leu Gly Phe His Asn Asn Asn Ile Lys Ala Ile Pro Glu Lys Ala
210 215 220

Phe Met Gly Asn Pro Leu Leu Gln Thr Ile His Phe Tyr Asp Asn Pro
225 230 235 240

Ile Gln Phe Val Gly Arg Ser Ala Phe Gln Tyr Leu Pro Lys Leu His
245 250 255

Thr Leu Ser Leu Asn Gly Ala Met Asp Ile Gln Glu Phe Pro Asp Leu
260 265 270

Lys Gly Thr Thr Ser Leu Glu Ile Leu Thr Leu Thr Arg Ala Gly Ile
275 280 285

Arg Leu Leu Pro Ser Gly Met Cys Gln Gln Leu Pro Arg Leu Arg Val

290 295 300

Leu Glu Leu Ser His Asn Gln Ile Glu Glu Leu Pro Ser Leu His Arg
305 310 315 320

Cys Gln Lys Leu Glu Glu Ile Gly Leu Gln His Asn Arg Ile Trp Glu
325 330 335

Ile Gly Ala Asp Thr Phe Ser Gln Leu Ser Ser Leu Gln Ala Leu Asp
340 345 350

Leu Ser Trp Asn Ala Ile Arg Ser Ile His Pro Glu Ala Phe Ser Thr
355 360 365

Leu His Ser Leu Val Lys Leu Asp Leu Thr Asp Asn Gln Leu Thr Thr
370 375 380

Leu Pro Leu Ala Gly Leu Gly Gly Leu Met His Leu Lys Leu Lys Gly
385 390 395 400

Asn Leu Ala Leu Ser Gln Ala Phe Ser Lys Asp Ser Phe Pro Lys Leu
405 410 415

Arg Ile Leu Glu Val Pro Tyr Ala Tyr Gln Cys Cys Pro Tyr Gly Met
420 425 430

Cys Ala Ser Phe Phe Lys Ala Ser Gly Gln Trp Glu Ala Glu Asp Leu
435 440 445

His Leu Asp Asp Glu Glu Ser Ser Lys Arg Pro Leu Gly Leu Leu Ala
450 455 460

Arg Gln Ala Glu Asn His Tyr Asp Gln Asp Leu Asp Glu Leu Gln Leu
465 470 475 480

Glu Met Glu Asp Ser Lys Pro His Pro Ser Val Gln Cys Ser Pro Thr
485 490 495

Pro Gly Pro Phe Lys Pro Cys Glu Tyr Leu Phe Glu Ser Trp Gly Ile
500 505 510

Arg Leu Ala Val Trp Ala Ile Val Leu Leu Ser Val Leu Cys Asn Gly
515 520 525

Leu Val Leu Leu Thr Val Phe Ala Gly Gly Pro Val Pro Leu Pro Pro
530 535 540

```
Val Lys Phe Val Val Gly Ala Ile Ala Gly Ala Asn Thr Leu Thr Gly
545                 550                 555                 560

Ile Ser Cys Gly Leu Leu Ala Ser Val Asp Ala Leu Thr Phe Gly Gln
                565                 570                 575

Phe Ser Glu Tyr Gly Ala Arg Trp Glu Thr Gly Leu Gly Cys Arg Ala
            580                 585                 590

Thr Gly Phe Leu Ala Val Leu Gly Ser Glu Ala Ser Val Leu Leu Leu
        595                 600                 605

Thr Leu Ala Ala Val Gln Cys Ser Val Ser Val Ser Cys Val Arg Ala
    610                 615                 620

Tyr Gly Lys Ser Pro Ser Leu Gly Ser Val Arg Ala Gly Val Leu Gly
625                 630                 635                 640

Cys Leu Ala Leu Ala Gly Leu Ala Ala Ala Leu Pro Leu Ala Ser Val
                645                 650                 655

Gly Glu Tyr Gly Ala Ser Pro Leu Cys Leu Pro Tyr Ala Pro Pro Glu
            660                 665                 670

Gly Gln Pro Ala Ala Leu Gly Phe Thr Val Ala Leu Val Met Met Asn
        675                 680                 685

Ser Phe Cys Phe Leu Val Val Ala Gly Ala Tyr Ile Lys Leu Tyr Cys
    690                 695                 700

Asp Leu Pro Arg Gly Asp Phe Glu Ala Val Trp Asp Cys Ala Met Val
705                 710                 715                 720

Arg His Val Ala Trp Leu Ile Phe Ala Asp Gly Leu Leu Tyr Cys Pro
                725                 730                 735

Val Ala Phe Leu Ser Phe Ala Ser Met Leu Gly Leu Phe Pro Val Thr
            740                 745                 750

Pro Glu Ala Val Lys Ser Val Leu Leu Val Val Leu Pro Leu Pro Ala
        755                 760                 765

Cys Leu Asn Pro Leu Leu Tyr Leu Leu Phe Asn Pro His Phe Arg Asp
    770                 775                 780
```

Asp Leu Arg Arg Leu Arg Pro Arg Ala Gly Asp Ser Gly Pro Leu Ala
785 790 795 800

Tyr Ala Ala Ala Gly Glu Leu Glu Lys Ser Ser Cys Asp Ser Thr Gln
805 810 815

Ala Leu Val Ala Phe Ser Asp Val Asp Leu Ile Leu Glu Ala Ser Glu
820 825 830

Ala Gly Arg Pro Pro Gly Leu Glu Thr Tyr Gly Phe Pro Ser Val Thr
835 840 845

Leu Ile Ser Cys Gln Gln Pro Gly Ala Pro Arg Leu Glu Gly Ser His
850 855 860

Cys Val Glu Pro Glu Gly Asn His Phe Gly Asn Pro Gln Pro Ser Met
865 870 875 880

Asp Gly Glu Leu Leu Leu Arg Ala Glu Gly Ser Thr Pro Ala Gly Gly
885 890 895

Gly Leu Ser Gly Gly Gly Gly Phe Gln Pro Ser Gly Leu Ala Phe Ala
900 905 910

Ser His Val
915

<210> 102
<211> 647
<212> PRT
<213> Homo Sapiens

<400> 102

Met Ala Ser Leu Val Ser Leu Glu Leu Gly Leu Leu Leu Ala Val Leu
1 5 10 15

Val Val Thr Ala Thr Ala Ser Pro Pro Ala Gly Leu Leu Ser Leu Leu
20 25 30

Thr Ser Gly Gln Gly Ala Leu Asp Gln Glu Ala Leu Gly Gly Leu Leu
35 40 45

Asn Thr Leu Ala Asp Arg Val His Cys Thr Asn Gly Pro Cys Gly Lys
50 55 60

Cys Leu Ser Val Glu Asp Ala Leu Gly Leu Gly Glu Pro Glu Gly Ser

65 70 75 80

Gly Leu Pro Pro Gly Pro Val Leu Glu Ala Arg Tyr Val Ala Arg Leu
85 90 95

Ser Ala Ala Ala Val Leu Tyr Leu Ser Asn Pro Glu Gly Thr Cys Glu
100 105 110

Asp Thr Arg Ala Gly Leu Trp Ala Ser His Ala Asp His Leu Leu Ala
115 120 125

Leu Leu Glu Ser Pro Lys Ala Leu Thr Pro Gly Leu Ser Trp Leu Leu
130 135 140

Gln Arg Met Gln Ala Arg Ala Ala Gly Gln Thr Pro Lys Thr Ala Cys
145 150 155 160

Val Asp Ile Pro Gln Leu Leu Glu Glu Ala Val Gly Ala Gly Ala Pro
165 170 175

Gly Ser Ala Gly Gly Val Leu Ala Ala Leu Leu Asp His Val Arg Ser
180 185 190

Gly Ser Cys Phe His Ala Leu Pro Ser Pro Gln Tyr Phe Val Asp Phe
195 200 205

Val Phe Gln Gln His Ser Ser Glu Val Pro Met Thr Leu Ala Glu Leu
210 215 220

Ser Ala Leu Met Gln Arg Leu Gly Val Gly Arg Glu Ala His Ser Asp
225 230 235 240

His Ser His Arg His Arg Gly Ala Ser Ser Arg Asp Pro Val Pro Leu
245 250 255

Ile Ser Ser Ser Asn Ser Ser Ser Val Trp Asp Thr Val Cys Leu Ser
260 265 270

Ala Arg Asp Val Met Ala Ala Tyr Gly Leu Ser Glu Gln Ala Gly Val
275 280 285

Thr Pro Glu Ala Trp Ala Gln Leu Ser Pro Ala Leu Leu Gln Gln Gln
290 295 300

Leu Ser Gly Ala Cys Thr Ser Gln Ser Arg Pro Pro Val Gln Asp Gln
305 310 315 320

Leu Ser Gln Ser Glu Arg Tyr Leu Tyr Gly Ser Leu Ala Thr Leu Leu
325 330 335

Ile Cys Leu Cys Ala Val Phe Gly Leu Leu Leu Leu Thr Cys Thr Gly
340 345 350

Cys Arg Gly Val Ala His Tyr Ile Leu Gln Thr Phe Leu Ser Leu Ala
355 360 365

Val Gly Ala Leu Thr Gly Asp Ala Val Leu His Leu Thr Pro Lys Val
370 375 380

Leu Gly Leu His Thr His Ser Glu Glu Gly Leu Ser Pro Gln Pro Thr
385 390 395 400

Trp Arg Leu Leu Ala Met Leu Ala Gly Leu Tyr Ala Phe Phe Leu Phe
405 410 415

Glu Asn Leu Phe Asn Leu Leu Leu Pro Arg Asp Pro Glu Asp Leu Glu
420 425 430

Asp Gly Pro Cys Gly His Ser Ser His Ser His Gly Gly His Ser His
435 440 445

Gly Val Ser Leu Gln Leu Ala Pro Ser Glu Leu Arg Gln Pro Lys Pro
450 455 460

Pro His Glu Gly Ser Arg Ala Asp Leu Val Ala Glu Glu Ser Pro Glu
465 470 475 480

Leu Leu Asn Pro Glu Pro Arg Arg Leu Ser Pro Glu Leu Arg Leu Leu
485 490 495

Pro Tyr Met Ile Thr Leu Gly Asp Ala Val His Asn Phe Ala Asp Gly
500 505 510

Leu Ala Val Gly Ala Ala Phe Ala Ser Ser Trp Lys Thr Gly Leu Ala
515 520 525

Thr Ser Leu Ala Val Phe Cys His Glu Leu Pro His Glu Leu Gly Asp
530 535 540

Phe Ala Ala Leu Leu His Ala Gly Leu Ser Val Arg Gln Ala Leu Leu
545 550 555 560

Leu Asn Leu Ala Ser Ala Leu Thr Ala Phe Ala Gly Leu Tyr Val Ala
565 570 575

Leu Ala Val Gly Val Ser Glu Glu Ser Glu Ala Trp Ile Leu Ala Val
580 585 590

Ala Thr Gly Leu Phe Leu Tyr Val Ala Leu Cys Asp Met Leu Pro Ala
595 600 605

Met Leu Lys Val Arg Asp Pro Arg Pro Trp Leu Leu Phe Leu Leu His
610 615 620

Asn Val Gly Leu Leu Gly Gly Trp Thr Val Leu Leu Leu Leu Ser Leu
625 630 635 640

Tyr Glu Asp Asp Ile Thr Phe
645

<210> 103
<211> 522
<212> PRT
<213> Homo Sapiens

<400> 103

Met Asp Phe Leu Leu Leu Gly Leu Cys Leu Tyr Trp Leu Leu Arg Arg
1 5 10 15

Pro Ser Gly Val Val Leu Cys Leu Leu Gly Ala Cys Phe Gln Met Leu
20 25 30

Pro Ala Ala Pro Ser Gly Cys Pro Gln Leu Cys Arg Cys Glu Gly Arg
35 40 45

Leu Leu Tyr Cys Glu Ala Leu Asn Leu Thr Glu Ala Pro His Asn Leu
50 55 60

Ser Gly Leu Leu Gly Leu Ser Leu Arg Tyr Asn Ser Leu Ser Glu Leu
65 70 75 80

Arg Ala Gly Gln Phe Thr Gly Leu Met Gln Leu Thr Trp Leu Tyr Leu
85 90 95

Asp His Asn His Ile Cys Ser Val Gln Gly Asp Ala Phe Gln Lys Leu
100 105 110

Arg Arg Val Lys Glu Leu Thr Leu Ser Ser Asn Gln Ile Thr Gln Leu

115 120 125

Pro Asn Thr Thr Phe Arg Pro Met Pro Asn Leu Arg Ser Val Asp Leu
130 135 140

Ser Tyr Asn Lys Leu Gln Ala Leu Ala Pro Asp Leu Phe His Gly Leu
145 150 155 160

Arg Lys Leu Thr Thr Leu His Met Arg Ala Asn Ala Ile Gln Phe Val
165 170 175

Pro Val Arg Ile Phe Gln Asp Cys Arg Ser Leu Lys Phe Leu Asp Ile
180 185 190

Gly Tyr Asn Gln Leu Lys Ser Leu Ala Arg Asn Ser Phe Ala Gly Leu
195 200 205

Phe Lys Leu Thr Glu Leu His Leu Glu His Asn Asp Leu Val Lys Val
210 215 220

Asn Phe Ala His Phe Pro Arg Leu Ile Ser Leu His Ser Leu Cys Leu
225 230 235 240

Arg Arg Asn Lys Val Ala Ile Val Val Ser Ser Leu Asp Trp Val Trp
245 250 255

Asn Leu Glu Lys Met Asp Leu Ser Gly Asn Glu Ile Glu Tyr Met Glu
260 265 270

Pro His Val Phe Glu Thr Val Pro His Leu Gln Ser Leu Gln Leu Asp
275 280 285

Ser Asn Arg Leu Thr Tyr Ile Glu Pro Arg Ile Leu Asn Ser Trp Lys
290 295 300

Ser Leu Thr Ser Ile Thr Leu Ala Gly Asn Leu Trp Asp Cys Gly Arg
305 310 315 320

Asn Val Cys Ala Leu Ala Ser Trp Leu Asn Asn Phe Gln Gly Arg Tyr
325 330 335

Asp Gly Asn Leu Gln Cys Ala Ser Pro Glu Tyr Ala Gln Gly Glu Asp
340 345 350

Val Leu Asp Ala Val Tyr Ala Phe His Leu Cys Glu Asp Gly Ala Glu
355 360 365

```
Pro Thr Ser Gly His Leu Leu Ser Ala Val Thr Asn Arg Ser Asp Leu
    370                 375                 380

Gly Pro Pro Ala Ser Ser Ala Thr Thr Leu Ala Asp Gly Gly Glu Gly
385                 390                 395                 400

Gln His Asp Gly Thr Phe Glu Pro Ala Thr Val Ala Leu Pro Gly Gly
                405                 410                 415

Glu His Ala Glu Asn Ala Val Gln Ile His Lys Val Val Thr Gly Thr
            420                 425                 430

Met Ala Leu Ile Phe Ser Phe Leu Ile Val Val Leu Val Leu Tyr Val
        435                 440                 445

Ser Trp Lys Cys Phe Pro Ala Ser Leu Arg Gln Leu Arg Gln Cys Phe
    450                 455                 460

Val Thr Gln Arg Arg Lys Gln Lys Gln Lys Gln Thr Met His Gln Met
465                 470                 475                 480

Ala Ala Met Ser Ala Gln Glu Tyr Tyr Val Asp Tyr Lys Pro Asn His
                485                 490                 495

Ile Glu Gly Ala Leu Val Thr Ile Asn Glu Tyr Gly Ser Cys Thr Cys
            500                 505                 510

His Gln Gln Pro Ala Arg Glu Cys Glu Val
        515                 520
```

<210> 104
<211> 375
<212> PRT
<213> Homo Sapiens

<400> 104

```
Met Ala Asn Ala Ser Glu Pro Gly Gly Ser Gly Gly Gly Glu Ala Ala
1               5                   10                  15

Ala Leu Gly Leu Lys Leu Ala Thr Leu Ser Leu Leu Leu Cys Val Ser
            20                  25                  30

Leu Ala Gly Asn Val Leu Phe Ala Leu Leu Ile Val Arg Glu Arg Ser
        35                  40                  45
```

```
Leu His Arg Ala Pro Tyr Tyr Leu Leu Leu Asp Leu Cys Leu Ala Asp
    50                  55                  60

Gly Leu Arg Ala Leu Ala Cys Leu Pro Ala Val Met Leu Ala Ala Arg
65                  70                  75                  80

Arg Ala Ala Ala Ala Ala Gly Ala Pro Pro Gly Ala Leu Gly Cys Lys
                85                  90                  95

Leu Leu Ala Phe Leu Ala Ala Leu Phe Cys Phe His Ala Ala Phe Leu
            100                 105                 110

Leu Leu Gly Val Gly Val Thr Arg Tyr Leu Ala Ile Ala His His Arg
        115                 120                 125

Phe Tyr Ala Glu Arg Leu Ala Gly Trp Pro Cys Ala Ala Met Leu Val
    130                 135                 140

Cys Ala Ala Trp Ala Leu Ala Leu Ala Ala Ala Phe Pro Pro Val Leu
145                 150                 155                 160

Asp Gly Gly Gly Asp Asp Glu Asp Ala Pro Cys Ala Leu Glu Gln Arg
                165                 170                 175

Pro Asp Gly Ala Pro Gly Ala Leu Gly Phe Leu Leu Leu Leu Ala Val
            180                 185                 190

Val Val Gly Ala Thr His Leu Val Tyr Leu Arg Leu Leu Phe Phe Ile
        195                 200                 205

His Asp Arg Arg Lys Met Arg Pro Ala Arg Leu Val Pro Ala Val Ser
    210                 215                 220

His Asp Trp Thr Phe His Gly Pro Gly Ala Thr Gly Gln Ala Ala Ala
225                 230                 235                 240

Asn Trp Thr Ala Gly Phe Gly Arg Gly Pro Thr Pro Pro Ala Leu Val
                245                 250                 255

Gly Ile Arg Pro Ala Gly Pro Gly Arg Gly Ala Arg Arg Leu Leu Val
            260                 265                 270

Leu Glu Glu Phe Lys Thr Glu Lys Arg Leu Cys Lys Met Phe Tyr Ala
        275                 280                 285

Val Thr Leu Leu Phe Leu Leu Leu Trp Gly Pro Tyr Val Val Ala Ser
```

290 295 300

Tyr Leu Arg Val Leu Val Arg Pro Gly Ala Val Pro Gln Ala Tyr Leu
305 310 315 320

Thr Ala Ser Val Trp Leu Thr Phe Ala Gln Ala Gly Ile Asn Pro Val
325 330 335

Val Cys Phe Leu Phe Asn Arg Glu Leu Arg Asp Cys Phe Arg Ala Gln
340 345 350

Phe Pro Cys Cys Gln Ser Pro Arg Thr Thr Gln Ala Thr His Pro Cys
355 360 365

Asp Leu Lys Gly Ile Gly Leu
370 375

<210> 105
<211> 349
<212> PRT
<213> Homo Sapiens

<400> 105

Met Asn Arg Lys Ala Arg Arg Cys Leu Gly His Leu Phe Leu Ser Leu
1 5 10 15

Gly Met Val Tyr Leu Arg Ile Gly Gly Phe Ser Ser Val Val Ala Leu
20 25 30

Gly Ala Ser Ile Ile Cys Asn Lys Ile Pro Gly Leu Ala Pro Arg Gln
35 40 45

Arg Ala Ile Cys Gln Ser Arg Pro Asp Ala Ile Ile Val Ile Gly Glu
50 55 60

Gly Ser Gln Met Gly Leu Asp Glu Cys Gln Phe Gln Phe Arg Asn Gly
65 70 75 80

Arg Trp Asn Cys Ser Ala Leu Gly Glu Arg Thr Val Phe Gly Lys Glu
85 90 95

Leu Lys Val Gly Ser Arg Glu Ala Ala Phe Thr Tyr Ala Ile Ile Ala
100 105 110

Ala Gly Val Ala His Ala Ile Thr Ala Ala Cys Thr Gln Gly Asn Leu
115 120 125

```
Ser Asp Cys Gly Cys Asp Lys Glu Lys Gln Gly Gln Tyr His Arg Asp
    130                 135                 140

Glu Gly Trp Lys Trp Gly Gly Cys Ser Ala Asp Ile Arg Tyr Gly Ile
145                 150                 155                 160

Gly Phe Ala Lys Val Phe Val Asp Ala Arg Glu Ile Lys Gln Asn Ala
                165                 170                 175

Arg Thr Leu Met Asn Leu His Asn Asn Glu Ala Gly Arg Lys Ile Leu
            180                 185                 190

Glu Glu Asn Met Lys Leu Glu Cys Lys Cys His Gly Val Ser Gly Ser
        195                 200                 205

Cys Thr Thr Lys Thr Cys Trp Thr Thr Leu Pro Gln Phe Arg Glu Leu
    210                 215                 220

Gly Tyr Val Leu Lys Asp Lys Tyr Asn Glu Ala Val His Val Glu Pro
225                 230                 235                 240

Val Arg Ala Ser Arg Asn Lys Arg Pro Thr Phe Leu Lys Ile Lys Lys
                245                 250                 255

Pro Leu Ser Tyr Arg Lys Pro Met Asp Thr Asp Leu Val Tyr Ile Glu
            260                 265                 270

Lys Ser Pro Asn Tyr Cys Glu Glu Asp Pro Val Thr Gly Ser Val Gly
        275                 280                 285

Thr Gln Gly Arg Ala Cys Asn Lys Thr Ala Pro Gln Ala Ser Gly Cys
    290                 295                 300

Asp Leu Met Cys Cys Gly Arg Gly Tyr Asn Thr His Gln Tyr Ala Arg
305                 310                 315                 320

Val Trp Gln Cys Asn Cys Lys Phe His Trp Cys Cys Tyr Val Lys Cys
                325                 330                 335

Asn Thr Cys Ser Glu Arg Thr Glu Met Tyr Thr Cys Lys
            340                 345
```

<210> 106
<211> 694
<212> PRT
<213> Homo Sapiens

<400> 106

Met Glu Trp Gly Tyr Leu Leu Glu Val Thr Ser Leu Leu Ala Ala Leu
1 5 10 15

Ala Leu Leu Gln Arg Ser Ser Gly Ala Ala Ala Ala Ser Ala Lys Glu
20 25 30

Leu Ala Cys Gln Glu Ile Thr Val Pro Leu Cys Lys Gly Ile Gly Tyr
35 40 45

Asn Tyr Thr Tyr Met Pro Asn Gln Phe Asn His Asp Thr Gln Asp Glu
50 55 60

Ala Gly Leu Glu Val His Gln Phe Trp Pro Leu Val Glu Ile Gln Cys
65 70 75 80

Ser Pro Asp Leu Lys Phe Phe Leu Cys Ser Met Tyr Thr Pro Ile Cys
85 90 95

Leu Glu Asp Tyr Lys Lys Pro Leu Pro Pro Cys Arg Ser Val Cys Glu
100 105 110

Arg Ala Lys Ala Gly Cys Ala Pro Leu Met Arg Gln Tyr Gly Phe Ala
115 120 125

Trp Pro Asp Arg Met Arg Cys Asp Arg Leu Pro Glu Gln Gly Asn Pro
130 135 140

Asp Thr Leu Cys Met Asp Tyr Asn Arg Thr Asp Leu Thr Thr Ala Ala
145 150 155 160

Pro Ser Pro Pro Arg Arg Leu Pro Pro Pro Pro Pro Gly Glu Gln Pro
165 170 175

Pro Ser Gly Ser Gly His Gly Arg Pro Pro Gly Ala Arg Pro Pro His
180 185 190

Arg Gly Gly Gly Arg Gly Gly Gly Gly Gly Asp Ala Ala Ala Pro Pro
195 200 205

Ala Arg Gly Gly Gly Gly Gly Gly Lys Ala Arg Pro Pro Gly Gly Gly
210 215 220

Ala Ala Pro Cys Glu Pro Gly Cys Gln Cys Arg Ala Pro Met Val Ser
225 230 235 240

```
Val Ser Ser Glu Arg His Pro Leu Tyr Asn Arg Val Lys Thr Gly Gln
                245                 250                 255

Ile Ala Asn Cys Ala Leu Pro Cys His Asn Pro Phe Phe Ser Gln Asp
            260                 265                 270

Glu Arg Ala Phe Thr Val Phe Trp Ile Gly Leu Trp Ser Val Leu Cys
        275                 280                 285

Phe Val Ser Thr Phe Ala Thr Val Ser Thr Phe Leu Ile Asp Met Glu
    290                 295                 300

Arg Phe Lys Tyr Pro Glu Arg Pro Ile Ile Phe Leu Ser Ala Cys Tyr
305                 310                 315                 320

Leu Phe Val Ser Val Gly Tyr Leu Val Arg Leu Val Ala Gly His Glu
                325                 330                 335

Lys Val Ala Cys Ser Gly Gly Ala Pro Gly Ala Gly Gly Ala Gly Gly
            340                 345                 350

Ala Gly Gly Ala Ala Ala Gly Ala Gly Ala Ala Gly Ala Gly Ala Gly
        355                 360                 365

Gly Pro Gly Gly Arg Gly Glu Tyr Glu Glu Leu Gly Ala Val Glu Gln
    370                 375                 380

His Val Arg Tyr Glu Thr Thr Gly Pro Ala Leu Cys Thr Val Val Phe
385                 390                 395                 400

Leu Leu Val Tyr Phe Phe Gly Met Ala Ser Ser Ile Trp Trp Val Ile
                405                 410                 415

Leu Ser Leu Thr Trp Phe Leu Ala Ala Gly Met Lys Trp Gly Asn Glu
            420                 425                 430

Ala Ile Ala Gly Tyr Ser Gln Tyr Phe His Leu Ala Ala Trp Leu Val
        435                 440                 445

Pro Ser Val Lys Ser Ile Ala Val Leu Ala Leu Ser Ser Val Asp Gly
    450                 455                 460

Asp Pro Val Ala Gly Ile Cys Tyr Val Gly Asn Gln Ser Leu Asp Asn
465                 470                 475                 480
```

Leu Arg Gly Phe Val Leu Ala Pro Leu Val Ile Tyr Leu Phe Ile Gly
485 490 495

Thr Met Phe Leu Leu Ala Gly Phe Val Ser Leu Phe Arg Ile Arg Ser
500 505 510

Val Ile Lys Gln Gln Asp Gly Pro Thr Lys Thr His Lys Leu Glu Lys
515 520 525

Leu Met Ile Arg Leu Gly Leu Phe Thr Val Leu Tyr Thr Val Pro Ala
530 535 540

Ala Val Val Val Ala Cys Leu Phe Tyr Glu Gln His Asn Arg Pro Arg
545 550 555 560

Trp Glu Ala Thr His Asn Cys Pro Cys Leu Arg Asp Leu Gln Pro Asp
565 570 575

Gln Ala Arg Arg Pro Asp Tyr Ala Val Phe Met Leu Lys Tyr Phe Met
580 585 590

Cys Leu Val Val Gly Ile Thr Ser Gly Val Trp Val Trp Ser Gly Lys
595 600 605

Thr Leu Glu Ser Trp Arg Ser Leu Cys Thr Arg Cys Cys Trp Ala Ser
610 615 620

Lys Gly Ala Ala Val Gly Gly Gly Ala Gly Ala Thr Ala Ala Gly Gly
625 630 635 640

Gly Gly Gly Pro Gly Gly Gly Gly Gly Gly Gly Pro Gly Gly Gly Gly
645 650 655

Gly Pro Gly Gly Gly Gly Gly Ser Leu Tyr Ser Asp Val Ser Thr Gly
660 665 670

Leu Thr Trp Arg Ser Gly Thr Ala Ser Ser Val Ser Tyr Pro Lys Gln
675 680 685

Met Pro Leu Ser Gln Val
690

<210> 107
<211> 295
<212> PRT
<213> Homo Sapiens

<400> 107

Met Leu Gln Gly Pro Gly Ser Leu Leu Leu Leu Phe Leu Ala Ser His
1 5 10 15

Cys Cys Leu Gly Ser Ala Arg Gly Leu Phe Leu Phe Gly Gln Pro Asp
20 25 30

Phe Ser Tyr Lys Arg Ser Asn Cys Lys Pro Ile Pro Ala Asn Leu Gln
35 40 45

Leu Cys His Gly Ile Glu Tyr Gln Asn Met Arg Leu Pro Asn Leu Leu
50 55 60

Gly His Glu Thr Met Lys Glu Val Leu Glu Gln Ala Gly Ala Trp Ile
65 70 75 80

Pro Leu Val Met Lys Gln Cys His Pro Asp Thr Lys Lys Phe Leu Cys
85 90 95

Ser Leu Phe Ala Pro Val Cys Leu Asp Asp Leu Asp Glu Thr Ile Gln
100 105 110

Pro Cys His Ser Leu Cys Val Gln Val Lys Asp Arg Cys Ala Pro Val
115 120 125

Met Ser Ala Phe Gly Phe Pro Trp Pro Asp Met Leu Glu Cys Asp Arg
130 135 140

Phe Pro Gln Asp Asn Asp Leu Cys Ile Pro Leu Ala Ser Ser Asp His
145 150 155 160

Leu Leu Pro Ala Thr Glu Glu Ala Pro Lys Val Cys Glu Ala Cys Lys
165 170 175

Asn Lys Asn Asp Asp Asp Asn Asp Ile Met Glu Thr Leu Cys Lys Asn
180 185 190

Asp Phe Ala Leu Lys Ile Lys Val Lys Glu Ile Thr Tyr Ile Asn Arg
195 200 205

Asp Thr Lys Ile Ile Leu Glu Thr Lys Ser Lys Thr Ile Tyr Lys Leu
210 215 220

Asn Gly Val Ser Glu Arg Asp Leu Lys Lys Ser Val Leu Trp Leu Lys
225 230 235 240

```
Asp Ser Leu Gln Cys Thr Cys Glu Glu Met Asn Asp Ile Asn Ala Pro
                245                 250                 255

Tyr Leu Val Met Gly Gln Lys Gln Gly Gly Glu Leu Val Ile Thr Ser
            260                 265                 270

Val Lys Arg Trp Gln Lys Gly Gln Arg Glu Phe Lys Arg Ile Ser Arg
        275                 280                 285

Ser Ile Arg Lys Leu Gln Cys
    290                 295
```

<210> 108
<211> 328
<212> PRT
<213> Homo Sapiens

<400> 108

```
Met Gly Phe Trp Ile Leu Ala Ile Leu Thr Ile Leu Met Tyr Ser Thr
1               5                   10                  15

Ala Ala Lys Phe Ser Lys Gln Ser Trp Gly Leu Glu Asn Glu Ala Leu
            20                  25                  30

Ile Val Arg Cys Pro Arg Gln Gly Lys Pro Ser Tyr Thr Val Asp Trp
        35                  40                  45

Tyr Tyr Ser Gln Thr Asn Lys Ser Ile Pro Thr Gln Glu Arg Asn Arg
    50                  55                  60

Val Phe Ala Ser Gly Gln Leu Leu Lys Phe Leu Pro Ala Ala Val Ala
65                  70                  75                  80

Asp Ser Gly Ile Tyr Thr Cys Ile Val Arg Ser Pro Thr Phe Asn Arg
                85                  90                  95

Thr Gly Tyr Ala Asn Val Thr Ile Tyr Lys Lys Gln Ser Asp Cys Asn
            100                 105                 110

Val Pro Asp Tyr Leu Met Tyr Ser Thr Val Ser Gly Ser Glu Lys Asn
        115                 120                 125

Ser Lys Ile Tyr Cys Pro Thr Ile Asp Leu Tyr Asn Trp Thr Ala Pro
    130                 135                 140
```

Leu Glu Trp Phe Lys Asn Cys Gln Ala Leu Gln Gly Ser Arg Tyr Arg
145 150 155 160

Ala His Lys Ser Phe Leu Val Ile Asp Asn Val Met Thr Glu Asp Ala
165 170 175

Gly Asp Tyr Thr Cys Lys Phe Ile His Asn Glu Asn Gly Ala Asn Tyr
180 185 190

Ser Val Thr Ala Thr Arg Ser Phe Thr Val Lys Asp Glu Gln Gly Phe
195 200 205

Ser Leu Phe Pro Val Ile Gly Ala Pro Ala Gln Asn Glu Ile Lys Glu
210 215 220

Val Glu Ile Gly Lys Asn Ala Asn Leu Thr Cys Ser Ala Cys Phe Gly
225 230 235 240

Lys Gly Thr Gln Phe Leu Ala Ala Val Leu Trp Gln Leu Asn Gly Thr
245 250 255

Lys Ile Thr Asp Phe Gly Glu Pro Arg Ile Gln Gln Glu Glu Gly Gln
260 265 270

Asn Gln Ser Phe Ser Asn Gly Leu Ala Cys Leu Asp Met Val Leu Arg
275 280 285

Ile Ala Asp Val Lys Glu Glu Asp Leu Leu Leu Gln Tyr Asp Cys Leu
290 295 300

Ala Leu Asn Leu His Gly Leu Arg Arg His Thr Val Arg Leu Ser Arg
305 310 315 320

Lys Asn Pro Ser Lys Glu Cys Phe
325

<210> 109
<211> 89
<212> PRT
<213> Homo Sapiens

<400> 109

Met Lys Gly Leu Ala Ala Ala Leu Leu Val Leu Val Cys Thr Met Ala
1 5 10 15

Leu Cys Ser Cys Ala Gln Val Gly Thr Asn Lys Glu Leu Cys Cys Leu
20 25 30

Val Tyr Thr Ser Trp Gln Ile Pro Gln Lys Phe Ile Val Asp Tyr Ser
35 40 45

Glu Thr Ser Pro Gln Cys Pro Lys Pro Gly Val Ile Leu Leu Thr Lys
50 55 60

Arg Gly Arg Gln Ile Cys Ala Asp Pro Asn Lys Lys Trp Val Gln Lys
65 70 75 80

Tyr Ile Ser Asp Leu Lys Leu Asn Ala
85

<210> 110
<211> 540
<212> PRT
<213> Homo Sapiens

<400> 110

Met Ala Thr Ala Pro Gly Pro Ala Gly Ile Ala Met Gly Ser Val Gly
1 5 10 15

Ser Leu Leu Glu Arg Gln Asp Phe Ser Pro Glu Glu Leu Arg Ala Ala
20 25 30

Leu Ala Gly Ser Arg Gly Ser Arg Gln Pro Asp Gly Leu Leu Arg Lys
35 40 45

Gly Leu Gly Gln Arg Glu Phe Leu Ser Tyr Leu His Leu Pro Lys Lys
50 55 60

Asp Ser Lys Ser Thr Lys Asn Thr Lys Arg Ala Pro Arg Asn Glu Pro
65 70 75 80

Ala Asp Tyr Ala Thr Leu Tyr Tyr Arg Glu His Ser Arg Ala Gly Asp
85 90 95

Phe Ser Lys Thr Ser Leu Pro Glu Arg Gly Arg Phe Asp Lys Cys Arg
100 105 110

Ile Arg Pro Ser Val Phe Lys Pro Thr Ala Gly Asn Gly Lys Gly Phe
115 120 125

Leu Ser Met Gln Ser Leu Ala Ser His Lys Gly Gln Lys Leu Trp Arg
130 135 140

```
Ser Asn Gly Ser Leu His Thr Leu Ala Cys His Pro Pro Leu Ser Pro
145                 150                 155                 160

Gly Pro Arg Ala Ser Gln Ala Arg Ala Gln Leu Leu His Ala Leu Ser
                165                 170                 175

Leu Asp Glu Gly Gly Pro Glu Pro Glu Pro Ser Leu Ser Asp Ser Ser
            180                 185                 190

Ser Gly Gly Ser Phe Gly Arg Ser Pro Gly Thr Gly Pro Ser Pro Phe
        195                 200                 205

Ser Ser Ser Leu Gly His Leu Asn His Leu Gly Gly Ser Leu Asp Arg
    210                 215                 220

Ala Ser Gln Gly Pro Lys Glu Ala Gly Pro Pro Ala Val Leu Ser Cys
225                 230                 235                 240

Leu Pro Glu Pro Pro Pro Pro Tyr Glu Phe Ser Cys Ser Ser Ala Glu
                245                 250                 255

Glu Met Gly Ala Val Leu Pro Glu Thr Cys Glu Glu Leu Lys Arg Gly
            260                 265                 270

Leu Gly Asp Glu Asp Gly Ser Asn Pro Phe Thr Gln Val Leu Glu Glu
        275                 280                 285

Arg Gln Arg Leu Trp Leu Ala Glu Leu Lys Arg Leu Tyr Val Glu Arg
    290                 295                 300

Leu His Glu Val Thr Gln Lys Ala Glu Arg Ser Glu Arg Asn Leu Gln
305                 310                 315                 320

Leu Gln Leu Phe Met Ala Gln Gln Glu Gln Arg Arg Leu Arg Lys Glu
                325                 330                 335

Leu Arg Ala Gln Gln Gly Leu Ala Pro Glu Pro Arg Ala Pro Gly Thr
            340                 345                 350

Leu Pro Glu Ala Asp Pro Ser Ala Arg Pro Glu Glu Glu Ala Arg Trp
        355                 360                 365

Glu Val Cys Gln Lys Thr Ala Glu Ile Ser Leu Leu Lys Gln Gln Leu
    370                 375                 380

Arg Glu Ala Gln Ala Glu Leu Ala Gln Lys Leu Ala Glu Ile Phe Ser
```

385 390 395 400

Leu Lys Thr Gln Leu Arg Gly Ser Arg Ala Gln Ala Gln Ala Gln Asp
405 410 415

Ala Glu Leu Val Arg Leu Arg Glu Ala Val Arg Ser Leu Gln Glu Gln
420 425 430

Ala Pro Arg Glu Glu Ala Pro Gly Ser Cys Glu Thr Asp Asp Cys Lys
435 440 445

Ser Arg Gly Leu Leu Gly Glu Ala Gly Gly Ser Glu Ala Arg Asp Ser
450 455 460

Ala Glu Gln Leu Arg Ala Glu Leu Leu Gln Glu Arg Leu Arg Gly Gln
465 470 475 480

Glu Gln Ala Leu Arg Phe Glu Gln Glu Arg Arg Thr Trp Gln Glu Glu
485 490 495

Lys Glu Arg Val Leu Arg Tyr Gln Arg Glu Ile Gln Gly Gly Tyr Met
500 505 510

Asp Met Tyr Arg Arg Asn Gln Ala Leu Glu Gln Glu Leu Arg Ala Leu
515 520 525

Arg Glu Pro Pro Thr Pro Trp Ser Pro Arg Leu Glu
530 535 540

<210> 111
<211> 673
<212> PRT
<213> Homo Sapiens

<400> 111

Met Pro Gly Gln Lys Phe Phe Leu Glu Val Leu Cys Cys Pro Ser Lys
1 5 10 15

Asn Trp Arg Ser Ser Ala Ala Glu Arg Val Pro Pro Ser Pro Ile Arg
20 25 30

Leu Arg Arg Arg Arg Pro Pro Ala Phe Ser Arg Arg Leu Pro Leu Arg
35 40 45

Arg Ser Asp Pro Ala Arg Ser Pro Gly Pro Ser Arg Arg Leu Ala Gly
50 55 60

```
Gly Phe Lys Ser Ala Arg Gly Ser Cys Asp Ala Gln Gly Leu Arg Ser
65                  70                  75                  80

Arg Gly Pro Ala Ser Ala Ser Pro Pro Trp Ala Ala Val Ser Ser Ile
                85                  90                  95

Ser Thr Lys Asp Trp Ser Glu Ser Asn Ser Ser Pro Cys Ser Glu Ile
            100                 105                 110

Pro Val Leu Pro Ala Asn Leu Gly Asp Trp Arg Gly Ile Trp Trp Gly
        115                 120                 125

Thr Trp Gln Glu Ala Pro Gly Pro Ala Gly Ile Ala Met Gly Ser Val
    130                 135                 140

Gly Ser Leu Leu Glu Arg Gln Asp Phe Ser Pro Glu Glu Leu Arg Ala
145                 150                 155                 160

Ala Leu Ala Gly Ser Arg Gly Ser Arg Gln Pro Asp Gly Leu Leu Arg
                165                 170                 175

Lys Gly Leu Gly Gln Arg Glu Phe Leu Ser Tyr Leu His Leu Pro Lys
            180                 185                 190

Lys Asp Ser Lys Ser Thr Lys Asn Thr Lys Arg Ala Pro Arg Asn Glu
        195                 200                 205

Pro Ala Asp Tyr Ala Thr Leu Tyr Tyr Arg Glu His Ser Arg Ala Gly
    210                 215                 220

Asp Phe Ser Lys Thr Ser Leu Pro Glu Arg Gly Arg Phe Asp Lys Cys
225                 230                 235                 240

Arg Ile Arg Pro Ser Val Phe Lys Pro Thr Ala Gly Asn Gly Lys Gly
                245                 250                 255

Phe Leu Ser Met Gln Ser Leu Ala Ser His Lys Gly Gln Lys Leu Trp
            260                 265                 270

Arg Ser Asn Gly Ser Leu His Thr Leu Ala Cys His Pro Pro Leu Ser
        275                 280                 285

Pro Gly Pro Arg Ala Ser Gln Ala Arg Ala Gln Leu Leu His Ala Leu
    290                 295                 300
```

Ser Leu Asp Glu Gly Gly Pro Glu Pro Glu Pro Ser Leu Ser Asp Ser
305 310 315 320

Ser Ser Gly Gly Ser Phe Gly Arg Ser Pro Gly Thr Gly Pro Ser Pro
325 330 335

Phe Ser Ser Ser Leu Gly His Leu Asn His Leu Gly Gly Ser Leu Asp
340 345 350

Arg Ala Ser Gln Gly Pro Lys Glu Ala Gly Pro Pro Ala Val Leu Ser
355 360 365

Cys Leu Pro Glu Pro Pro Pro Pro Tyr Glu Phe Ser Cys Ser Ser Ala
370 375 380

Glu Glu Met Gly Ala Val Leu Pro Glu Thr Cys Glu Glu Leu Lys Arg
385 390 395 400

Gly Leu Gly Asp Glu Asp Gly Ser Asn Pro Phe Thr Gln Val Leu Glu
405 410 415

Glu Arg Gln Arg Leu Trp Leu Ala Glu Leu Lys Arg Leu Tyr Val Glu
420 425 430

Arg Leu His Glu Val Thr Gln Lys Ala Glu Arg Ser Glu Arg Asn Leu
435 440 445

Gln Leu Gln Leu Phe Met Ala Gln Gln Glu Gln Arg Arg Leu Arg Lys
450 455 460

Glu Leu Arg Ala Gln Gln Gly Leu Ala Pro Glu Pro Arg Ala Pro Gly
465 470 475 480

Thr Leu Pro Glu Ala Asp Pro Ser Ala Arg Pro Glu Glu Glu Ala Arg
485 490 495

Trp Glu Val Cys Gln Lys Thr Ala Glu Ile Ser Leu Leu Lys Gln Gln
500 505 510

Leu Arg Glu Ala Gln Ala Glu Leu Ala Gln Lys Leu Ala Glu Ile Phe
515 520 525

Ser Leu Lys Thr Gln Leu Arg Gly Ser Arg Ala Gln Ala Gln Ala Gln
530 535 540

Asp Ala Glu Leu Val Arg Leu Arg Glu Ala Val Arg Ser Leu Gln Glu

545 550 555 560

Gln Ala Pro Arg Glu Glu Ala Pro Gly Ser Cys Glu Thr Asp Asp Cys
565 570 575

Lys Ser Arg Gly Leu Leu Gly Glu Ala Gly Gly Ser Glu Ala Arg Asp
580 585 590

Ser Ala Glu Gln Leu Arg Ala Glu Leu Leu Gln Glu Arg Leu Arg Gly
595 600 605

Gln Glu Gln Ala Leu Arg Phe Glu Gln Glu Arg Arg Thr Trp Gln Glu
610 615 620

Glu Lys Glu Arg Val Leu Arg Tyr Gln Arg Glu Ile Gln Gly Gly Tyr
625 630 635 640

Met Asp Met Tyr Arg Arg Asn Gln Ala Leu Glu Gln Glu Leu Arg Ala
645 650 655

Leu Arg Glu Pro Pro Thr Pro Trp Ser Pro Arg Leu Glu Ser Ser Lys
660 665 670

Ile

<210> 112
<211> 998
<212> PRT
<213> Homo Sapiens

<400> 112

Met Ala Arg Ala Arg Pro Pro Pro Pro Pro Ser Pro Pro Pro Gly Leu
1 5 10 15

Leu Pro Leu Leu Pro Pro Leu Leu Leu Leu Pro Leu Leu Leu Leu Pro
20 25 30

Ala Gly Cys Arg Ala Leu Glu Glu Thr Leu Met Asp Thr Lys Trp Val
35 40 45

Thr Ser Glu Leu Ala Trp Thr Ser His Pro Glu Ser Gly Trp Glu Glu
50 55 60

Val Ser Gly Tyr Asp Glu Ala Met Asn Pro Ile Arg Thr Tyr Gln Val
65 70 75 80

```
Cys Asn Val Arg Glu Ser Ser Gln Asn Asn Trp Leu Arg Thr Gly Phe
                85                  90                  95

Ile Trp Arg Arg Asp Val Gln Arg Val Tyr Val Glu Leu Lys Phe Thr
            100                 105                 110

Val Arg Asp Cys Asn Ser Ile Pro Asn Ile Pro Gly Ser Cys Lys Glu
        115                 120                 125

Thr Phe Asn Leu Phe Tyr Tyr Glu Ala Asp Ser Asp Val Ala Ser Ala
    130                 135                 140

Ser Ser Pro Phe Trp Met Glu Asn Pro Tyr Val Lys Val Asp Thr Ile
145                 150                 155                 160

Ala Pro Asp Glu Ser Phe Ser Arg Leu Asp Ala Gly Arg Val Asn Thr
                165                 170                 175

Lys Val Arg Ser Phe Gly Pro Leu Ser Lys Ala Gly Phe Tyr Leu Ala
            180                 185                 190

Phe Gln Asp Gln Gly Ala Cys Met Ser Leu Ile Ser Val Arg Ala Phe
        195                 200                 205

Tyr Lys Lys Cys Ala Ser Thr Thr Ala Gly Phe Ala Leu Phe Pro Glu
    210                 215                 220

Thr Leu Thr Gly Ala Glu Pro Thr Ser Leu Val Ile Ala Pro Gly Thr
225                 230                 235                 240

Cys Ile Pro Asn Ala Val Glu Val Ser Val Pro Leu Lys Leu Tyr Cys
                245                 250                 255

Asn Gly Asp Gly Glu Trp Met Val Pro Val Gly Ala Cys Thr Cys Ala
            260                 265                 270

Thr Gly His Glu Pro Ala Ala Lys Glu Ser Gln Cys Arg Pro Cys Pro
        275                 280                 285

Pro Gly Ser Tyr Lys Ala Lys Gln Gly Glu Gly Pro Cys Leu Pro Cys
    290                 295                 300

Pro Pro Asn Ser Arg Thr Thr Ser Pro Ala Ala Ser Ile Cys Thr Cys
305                 310                 315                 320
```

```
His Asn Asn Phe Tyr Arg Ala Asp Ser Asp Ser Ala Asp Ser Ala Cys
                325                 330                 335

Thr Thr Val Pro Ser Pro Pro Arg Gly Val Ile Ser Asn Val Asn Glu
            340                 345                 350

Thr Ser Leu Ile Leu Glu Trp Ser Glu Pro Arg Asp Leu Gly Gly Arg
        355                 360                 365

Asp Asp Leu Leu Tyr Asn Val Ile Cys Lys Lys Cys His Gly Ala Gly
    370                 375                 380

Gly Ala Ser Ala Cys Ser Arg Cys Asp Asp Asn Val Glu Phe Val Pro
385                 390                 395                 400

Arg Gln Leu Gly Leu Thr Glu Arg Arg Val His Ile Ser His Leu Leu
                405                 410                 415

Ala His Thr Arg Tyr Thr Phe Glu Val Gln Ala Val Asn Gly Val Ser
            420                 425                 430

Gly Lys Ser Pro Leu Pro Pro Arg Tyr Ala Ala Val Asn Ile Thr Thr
        435                 440                 445

Asn Gln Ala Ala Pro Ser Glu Val Pro Thr Leu Arg Leu His Ser Ser
    450                 455                 460

Ser Gly Ser Ser Leu Thr Leu Ser Trp Ala Pro Pro Glu Arg Pro Asn
465                 470                 475                 480

Gly Val Ile Leu Asp Tyr Glu Met Lys Tyr Phe Glu Lys Ser Glu Gly
                485                 490                 495

Ile Ala Ser Thr Val Thr Ser Gln Met Asn Ser Val Gln Leu Asp Gly
            500                 505                 510

Leu Arg Pro Asp Ala Arg Tyr Val Val Gln Val Arg Ala Arg Thr Val
        515                 520                 525

Ala Gly Tyr Gly Gln Tyr Ser Arg Pro Ala Glu Phe Glu Thr Thr Ser
    530                 535                 540

Glu Arg Gly Ser Gly Ala Gln Gln Leu Gln Glu Gln Leu Pro Leu Ile
545                 550                 555                 560

Val Gly Ser Ala Thr Ala Gly Leu Val Phe Val Val Ala Val Val Val
```

565 570 575

Ile Ala Ile Val Cys Leu Arg Lys Gln Arg His Gly Ser Asp Ser Glu
580 585 590

Tyr Thr Glu Lys Leu Gln Gln Tyr Ile Ala Pro Gly Met Lys Val Tyr
595 600 605

Ile Asp Pro Phe Thr Tyr Glu Asp Pro Asn Glu Ala Val Arg Glu Phe
610 615 620

Ala Lys Glu Ile Asp Val Ser Cys Val Lys Ile Glu Glu Val Ile Gly
625 630 635 640

Ala Gly Glu Phe Gly Glu Val Cys Arg Gly Arg Leu Lys Gln Pro Gly
645 650 655

Arg Arg Glu Val Phe Val Ala Ile Lys Thr Leu Lys Val Gly Tyr Thr
660 665 670

Glu Arg Gln Arg Arg Asp Phe Leu Ser Glu Ala Ser Ile Met Gly Gln
675 680 685

Phe Asp His Pro Asn Ile Ile Arg Leu Glu Gly Val Val Thr Lys Ser
690 695 700

Arg Pro Val Met Ile Leu Thr Glu Phe Met Glu Asn Cys Ala Leu Asp
705 710 715 720

Ser Phe Leu Arg Leu Asn Asp Gly Gln Phe Thr Val Ile Gln Leu Val
725 730 735

Gly Met Leu Arg Gly Ile Ala Ala Gly Met Lys Tyr Leu Ser Glu Met
740 745 750

Asn Tyr Val His Arg Asp Leu Ala Ala Arg Asn Ile Leu Val Asn Ser
755 760 765

Asn Leu Val Cys Lys Val Ser Asp Phe Gly Leu Ser Arg Phe Leu Glu
770 775 780

Asp Asp Pro Ser Asp Pro Thr Tyr Thr Ser Ser Leu Gly Gly Lys Ile
785 790 795 800

Pro Ile Arg Trp Thr Ala Pro Glu Ala Ile Ala Tyr Arg Lys Phe Thr
805 810 815

Ser Ala Ser Asp Val Trp Ser Tyr Gly Ile Val Met Trp Glu Val Met
820 825 830

Ser Tyr Gly Glu Arg Pro Tyr Trp Asp Met Ser Asn Gln Asp Val Ile
835 840 845

Asn Ala Val Glu Gln Asp Tyr Arg Leu Pro Pro Pro Met Asp Cys Pro
850 855 860

Thr Ala Leu His Gln Leu Met Leu Asp Cys Trp Val Arg Asp Arg Asn
865 870 875 880

Leu Arg Pro Lys Phe Ser Gln Ile Val Asn Thr Leu Asp Lys Leu Ile
885 890 895

Arg Asn Ala Ala Ser Leu Lys Val Ile Ala Ser Ala Gln Ser Gly Met
900 905 910

Ser Gln Pro Leu Leu Asp Arg Thr Val Pro Asp Tyr Thr Thr Phe Thr
915 920 925

Thr Val Gly Asp Trp Leu Asp Ala Ile Lys Met Gly Arg Tyr Lys Glu
930 935 940

Ser Phe Val Ser Ala Gly Phe Ala Ser Phe Asp Leu Val Ala Gln Met
945 950 955 960

Thr Ala Glu Asp Leu Leu Arg Ile Gly Val Thr Leu Ala Gly His Gln
965 970 975

Lys Lys Ile Leu Ser Ser Ile Gln Asp Met Arg Leu Gln Met Asn Gln
980 985 990

Thr Leu Pro Val Gln Val
995

<210> 113
<211> 413
<212> PRT
<213> Homo Sapiens

<400> 113

Met Gly Gly Thr Thr Leu Ala Trp Ser Met Ala Arg Asp Ser Ala Gly
1 5 10 15

Leu Val Ala Gly Asn Leu Asp Leu Ser Glu Lys His Asp Pro Arg Pro
20 25 30

Pro Pro Leu Leu His Pro Pro Gly Pro Thr Ala Val Leu Ala Gly Asp
35 40 45

Gly Ser Phe Arg Lys Cys Ala Glu Lys Ser Thr Phe Pro Cys Gln Ala
50 55 60

Thr Ala Arg Glu Leu Thr Pro Leu Phe Glu Pro Cys Gln Pro Pro His
65 70 75 80

Leu Val Gly Arg Val Lys Gly Arg Glu Val Asn Thr Ala Pro Thr Pro
85 90 95

Leu Pro Cys Arg Pro Ser Gly Arg Pro Val Ala Gly Gly Gly Gly Asp
100 105 110

Gly Pro Gly Gly Pro Glu Pro Gly Trp Val Asp Pro Arg Thr Trp Leu
115 120 125

Ser Phe Gln Gly Pro Pro Gly Gly Pro Gly Ile Gly Pro Gly Val Gly
130 135 140

Pro Gly Ser Glu Val Trp Gly Ile Pro Pro Cys Pro Pro Pro Tyr Glu
145 150 155 160

Phe Cys Gly Gly Met Ala Tyr Cys Gly Pro Gln Val Gly Val Gly Leu
165 170 175

Val Pro Gln Gly Gly Leu Glu Thr Ser Gln Pro Glu Gly Glu Ala Gly
180 185 190

Val Gly Val Glu Ser Asn Ser Asp Gly Ala Ser Pro Glu Pro Cys Thr
195 200 205

Val Thr Pro Gly Ala Val Lys Leu Glu Lys Glu Lys Leu Glu Gln Asn
210 215 220

Pro Glu Glu Ala Arg Lys Val Phe Ser Gln Thr Thr Ile Cys Arg Phe
225 230 235 240

Glu Ala Leu Gln Leu Ser Phe Lys Asn Met Cys Lys Leu Arg Pro Leu
245 250 255

Leu Gln Lys Trp Val Glu Glu Ala Asp Asn Asn Glu Asn Leu Gln Glu

260 265 270

Ile Cys Lys Ala Glu Thr Leu Val Gln Ala Arg Lys Arg Lys Arg Thr
275 280 285

Ser Ile Glu Asn Arg Val Arg Gly Asn Leu Glu Asn Leu Phe Leu Gln
290 295 300

Cys Pro Lys Pro Thr Leu Gln Gln Ile Ser His Ile Ala Gln Gln Leu
305 310 315 320

Gly Leu Glu Lys Asp Val Val Arg Val Trp Phe Cys Asn Arg Arg Gln
325 330 335

Lys Gly Lys Arg Ser Ser Ser Asp Tyr Ala Gln Arg Glu Asp Phe Glu
340 345 350

Ala Ala Gly Ser Pro Phe Ser Gly Gly Pro Val Ser Phe Pro Leu Ala
355 360 365

Pro Gly Pro His Phe Gly Thr Pro Gly Tyr Gly Ser Pro His Phe Thr
370 375 380

Ala Leu Tyr Ser Ser Val Pro Phe Pro Glu Gly Glu Ala Phe Pro Pro
385 390 395 400

Val Ser Val Thr Thr Leu Gly Ser Pro Met His Ser Asn
405 410

<210> 114
<211> 360
<212> PRT
<213> Homo Sapiens

<400> 114

Met Ala Gly His Leu Ala Ser Asp Phe Ala Phe Ser Pro Pro Pro Gly
1 5 10 15

Gly Gly Gly Asp Gly Pro Gly Gly Pro Glu Pro Gly Trp Val Asp Pro
20 25 30

Arg Thr Trp Leu Ser Phe Gln Gly Pro Pro Gly Gly Pro Gly Ile Gly
35 40 45

Pro Gly Val Gly Pro Gly Ser Glu Val Trp Gly Ile Pro Pro Cys Pro
50 55 60

```
Pro Pro Tyr Glu Phe Cys Gly Gly Met Ala Tyr Cys Gly Pro Gln Val
65                  70                  75                  80

Gly Val Gly Leu Val Pro Gln Gly Gly Leu Glu Thr Ser Gln Pro Glu
                85                  90                  95

Gly Glu Ala Gly Val Gly Val Glu Ser Asn Ser Asp Gly Ala Ser Pro
            100                 105                 110

Glu Pro Cys Thr Val Thr Pro Gly Ala Val Lys Leu Glu Lys Glu Lys
        115                 120                 125

Leu Glu Gln Asn Pro Glu Glu Ser Gln Asp Ile Lys Ala Leu Gln Lys
    130                 135                 140

Glu Leu Glu Gln Phe Ala Lys Leu Leu Lys Gln Lys Arg Ile Thr Leu
145                 150                 155                 160

Gly Tyr Thr Gln Ala Asp Val Gly Leu Thr Leu Gly Val Leu Phe Gly
                165                 170                 175

Lys Val Phe Ser Gln Thr Thr Ile Cys Arg Phe Glu Ala Leu Gln Leu
            180                 185                 190

Ser Phe Lys Asn Met Cys Lys Leu Arg Pro Leu Leu Gln Lys Trp Val
        195                 200                 205

Glu Glu Ala Asp Asn Asn Glu Asn Leu Gln Glu Ile Cys Lys Ala Glu
    210                 215                 220

Thr Leu Val Gln Ala Arg Lys Arg Lys Arg Thr Ser Ile Glu Asn Arg
225                 230                 235                 240

Val Arg Gly Asn Leu Glu Asn Leu Phe Leu Gln Cys Pro Lys Pro Thr
                245                 250                 255

Leu Gln Gln Ile Ser His Ile Ala Gln Gln Leu Gly Leu Glu Lys Asp
            260                 265                 270

Val Val Arg Val Trp Phe Cys Asn Arg Arg Gln Lys Gly Lys Arg Ser
        275                 280                 285

Ser Ser Asp Tyr Ala Gln Arg Glu Asp Phe Glu Ala Ala Gly Ser Pro
    290                 295                 300
```

Phe Ser Gly Gly Pro Val Ser Phe Pro Leu Ala Pro Gly Pro His Phe
305 310 315 320

Gly Thr Pro Gly Tyr Gly Ser Pro His Phe Thr Ala Leu Tyr Ser Ser
325 330 335

Val Pro Phe Pro Glu Gly Glu Ala Phe Pro Pro Val Ser Val Thr Thr
340 345 350

Leu Gly Ser Pro Met His Ser Asn
355 360

<210> 115
<211> 529
<212> PRT
<213> Homo Sapiens

<400> 115

Met Ser Val Lys Trp Thr Ser Val Ile Leu Leu Ile Gln Leu Ser Phe
1 5 10 15

Cys Phe Ser Ser Gly Asn Cys Gly Lys Val Leu Val Trp Ala Ala Glu
20 25 30

Tyr Ser His Trp Met Asn Ile Lys Thr Ile Leu Asp Glu Leu Ile Gln
35 40 45

Arg Gly His Glu Val Thr Val Leu Ala Ser Ser Ala Ser Ile Leu Phe
50 55 60

Asp Pro Asn Asn Ser Ser Ala Leu Lys Ile Glu Ile Tyr Pro Thr Ser
65 70 75 80

Leu Thr Lys Thr Glu Leu Glu Asn Phe Ile Met Gln Gln Ile Lys Arg
85 90 95

Trp Ser Asp Leu Pro Lys Asp Thr Phe Trp Leu Tyr Phe Ser Gln Val
100 105 110

Gln Glu Ile Met Ser Ile Phe Gly Asp Ile Thr Arg Lys Phe Cys Lys
115 120 125

Asp Val Val Ser Asn Lys Lys Phe Met Lys Lys Val Gln Glu Ser Arg
130 135 140

Phe Asp Val Ile Phe Ala Asp Ala Ile Phe Pro Cys Ser Glu Leu Leu
145 150 155 160

Ala Glu Leu Phe Asn Ile Pro Phe Val Tyr Ser Leu Ser Phe Ser Pro
165 170 175

Gly Tyr Thr Phe Glu Lys His Ser Gly Gly Phe Ile Phe Pro Pro Ser
180 185 190

Tyr Val Pro Val Val Met Ser Glu Leu Thr Asp Gln Met Thr Phe Met
195 200 205

Glu Arg Val Lys Asn Met Ile Tyr Val Leu Tyr Phe Asp Phe Trp Phe
210 215 220

Glu Ile Phe Asp Met Lys Lys Trp Asp Gln Phe Tyr Ser Glu Val Leu
225 230 235 240

Gly Arg Pro Thr Thr Leu Ser Glu Thr Met Gly Lys Ala Asp Val Trp
245 250 255

Leu Ile Arg Asn Ser Trp Asn Phe Gln Phe Pro His Pro Leu Leu Pro
260 265 270

Asn Val Asp Phe Val Gly Gly Leu His Cys Lys Pro Ala Lys Pro Leu
275 280 285

Pro Lys Glu Met Glu Asp Phe Val Gln Ser Ser Gly Glu Asn Gly Val
290 295 300

Val Val Phe Ser Leu Gly Ser Met Val Ser Asn Met Thr Glu Glu Arg
305 310 315 320

Ala Asn Val Ile Ala Ser Ala Leu Ala Gln Ile Pro Gln Lys Val Leu
325 330 335

Trp Arg Phe Asp Gly Asn Lys Pro Asp Thr Leu Gly Leu Asn Thr Arg
340 345 350

Leu Tyr Lys Trp Ile Pro Gln Asn Asp Leu Leu Gly His Pro Lys Thr
355 360 365

Arg Ala Phe Ile Thr His Gly Gly Ala Asn Gly Ile Tyr Glu Ala Ile
370 375 380

Tyr His Gly Ile Pro Met Val Gly Ile Pro Leu Phe Ala Asp Gln Pro
385 390 395 400

Asp Asn Ile Ala His Met Lys Ala Arg Gly Ala Ala Val Arg Val Asp
405 410 415

Phe Asn Thr Met Ser Ser Thr Asp Leu Leu Asn Ala Leu Lys Arg Val
420 425 430

Ile Asn Asp Pro Ser Tyr Lys Glu Asn Val Met Lys Leu Ser Arg Ile
435 440 445

Gln His Asp Gln Pro Val Lys Pro Leu Asp Arg Ala Val Phe Trp Ile
450 455 460

Glu Phe Val Met Arg His Lys Gly Ala Lys His Leu Arg Val Ala Ala
465 470 475 480

His Asp Leu Thr Trp Phe Gln Tyr His Ser Leu Asp Val Ile Gly Phe
485 490 495

Leu Leu Val Cys Val Ala Thr Val Ile Phe Ile Val Thr Lys Cys Cys
500 505 510

Leu Phe Cys Phe Trp Lys Phe Ala Arg Lys Ala Lys Lys Gly Lys Asn
515 520 525

Asp

<210> 116
<211> 2872
<212> PRT
<213> Homo Sapiens

<400> 116

Met Leu Gln Cys Thr Pro Ala Asn Met Val Glu Val His Lys Asp Lys
1 5 10 15

Glu Ser Ser Lys Gly His Thr Arg His Lys Val Glu Glu Ala Leu Ile
20 25 30

Asn Glu Glu Ala Ile Leu Asn Leu Met Glu Asn Ser Gln Thr Phe Gln
35 40 45

Pro Leu Thr Gln Arg Leu Ser Glu Ser Pro Val Phe Met Asp Ser Ser
50 55 60

Pro Asp Glu Ala Leu Val His Leu Leu Ala Gly Leu Glu Ser Asp Gly

65 70 75 80

Tyr Arg Gly Glu Arg Asn Arg Met Pro Ser Pro Cys Arg Ser Phe Gly
85 90 95

Asn Asn Lys Tyr Pro Gln Asn Ser Asp Asp Glu Glu Asn Glu Pro Gln
100 105 110

Ile Glu Lys Glu Glu Met Glu Leu Ser Leu Val Met Ser Gln Arg Trp
115 120 125

Asp Ser Asn Ile Glu Glu His Cys Ala Lys Lys Arg Ser Leu Cys Arg
130 135 140

Asn Thr His Arg Ser Ser Thr Glu Asp Asp Asp Ser Ser Ser Gly Glu
145 150 155 160

Glu Met Glu Trp Ser Asp Asn Ser Leu Leu Leu Ala Ser Leu Ser Ile
165 170 175

Pro Gln Leu Asp Gly Thr Ala Asp Glu Asn Ser Asp Asn Pro Leu Asn
180 185 190

Asn Glu Asn Ser Arg Thr His Ser Ser Val Ile Ala Thr Ser Lys Leu
195 200 205

Ser Val Lys Pro Ser Ile Phe His Lys Asp Ala Ala Thr Leu Glu Pro
210 215 220

Ser Ser Ser Ala Lys Ile Thr Phe Gln Cys Lys His Thr Ser Ala Leu
225 230 235 240

Ser Ser His Val Leu Asn Lys Glu Asp Leu Ile Glu Asp Leu Ser Gln
245 250 255

Thr Asn Lys Asn Thr Glu Lys Gly Leu Asp Asn Ser Val Thr Ser Phe
260 265 270

Thr Asn Glu Ser Thr Tyr Ser Met Lys Tyr Pro Gly Ser Leu Ser Ser
275 280 285

Thr Val His Ser Glu Asn Ser His Lys Glu Asn Ser Lys Lys Glu Ile
290 295 300

Leu Pro Val Ser Ser Cys Glu Ser Ser Ile Phe Asp Tyr Glu Glu Asp
305 310 315 320

```
Ile Pro Ser Val Thr Arg Gln Val Pro Ser Arg Lys Tyr Thr Asn Ile
                325                 330                 335

Arg Lys Ile Glu Lys Asp Ser Pro Phe Ile His Met His Arg His Pro
            340                 345                 350

Asn Glu Asn Thr Leu Gly Lys Asn Ser Phe Asn Phe Ser Asp Leu Asn
        355                 360                 365

His Ser Lys Asn Lys Val Ser Ser Glu Gly Asn Glu Lys Gly Asn Ser
    370                 375                 380

Thr Ala Leu Ser Ser Leu Phe Pro Ser Ser Phe Thr Glu Asn Cys Glu
385                 390                 395                 400

Leu Leu Ser Cys Ser Gly Glu Asn Arg Thr Met Val His Ser Leu Asn
                405                 410                 415

Ser Thr Ala Asp Glu Ser Gly Leu Asn Lys Leu Lys Ile Arg Tyr Glu
            420                 425                 430

Glu Phe Gln Glu His Lys Thr Glu Lys Pro Ser Leu Ser Gln Gln Ala
        435                 440                 445

Ala His Tyr Met Phe Phe Pro Ser Val Val Leu Ser Asn Cys Leu Thr
    450                 455                 460

Arg Pro Gln Lys Leu Ser Pro Val Thr Tyr Lys Leu Gln Pro Gly Asn
465                 470                 475                 480

Lys Pro Ser Arg Leu Lys Leu Asn Lys Arg Lys Leu Ala Gly His Gln
                485                 490                 495

Glu Thr Ser Thr Lys Ser Ser Glu Thr Gly Ser Thr Lys Asp Asn Phe
            500                 505                 510

Ile Gln Asn Asn Pro Cys Asn Ser Asn Pro Glu Lys Asp Asn Ala Leu
        515                 520                 525

Ala Ser Asp Leu Thr Lys Thr Thr Arg Gly Ala Phe Glu Asn Lys Thr
    530                 535                 540

Pro Thr Asp Gly Phe Ile Asp Cys His Phe Gly Asp Gly Thr Leu Glu
545                 550                 555                 560
```

```
Thr Glu Gln Ser Phe Gly Leu Tyr Gly Asn Lys Tyr Thr Leu Arg Ala
                565                 570                 575

Lys Arg Lys Val Asn Tyr Glu Thr Glu Asp Ser Glu Ser Ser Phe Val
            580                 585                 590

Thr His Asn Ser Lys Ile Ser Leu Pro His Pro Met Glu Ile Gly Glu
        595                 600                 605

Ser Leu Asp Gly Thr Leu Lys Ser Arg Lys Arg Arg Lys Met Ser Lys
    610                 615                 620

Lys Leu Pro Pro Val Ile Ile Lys Tyr Ile Ile Ile Asn Arg Phe Arg
625                 630                 635                 640

Gly Arg Lys Asn Met Leu Val Lys Leu Gly Lys Ile Asp Ser Lys Glu
                645                 650                 655

Lys Gln Val Ile Leu Thr Glu Glu Lys Met Glu Leu Tyr Lys Lys Leu
            660                 665                 670

Ala Pro Leu Lys Asp Phe Trp Pro Lys Val Pro Asp Ser Pro Ala Thr
        675                 680                 685

Lys Tyr Pro Ile Tyr Pro Leu Thr Pro Lys Lys Ser His Arg Arg Lys
    690                 695                 700

Ser Lys His Lys Ser Ala Lys Lys Lys Thr Gly Lys Gln Gln Arg Thr
705                 710                 715                 720

Asn Asn Glu Asn Ile Lys Arg Thr Leu Ser Phe Arg Lys Lys Arg Ser
                725                 730                 735

His Ala Ile Leu Ser Pro Pro Ser Pro Ser Tyr Asn Ala Glu Thr Glu
            740                 745                 750

Asp Cys Asp Leu Asn Tyr Ser Asp Val Met Ser Lys Leu Gly Phe Leu
        755                 760                 765

Ser Glu Arg Ser Thr Ser Pro Ile Asn Ser Ser Pro Pro Arg Cys Trp
    770                 775                 780

Ser Pro Thr Asp Pro Arg Ala Glu Glu Ile Met Ala Ala Ala Glu Lys
785                 790                 795                 800
```

Glu Ala Met Leu Phe Lys Gly Pro Asn Val Tyr Lys Lys Thr Val Asn
805 810 815

Ser Arg Ile Gly Lys Thr Ser Arg Ala Arg Ala Gln Ile Lys Lys Ser
820 825 830

Lys Ala Lys Leu Ala Asn Pro Ser Ile Val Thr Lys Lys Arg Asn Lys
835 840 845

Arg Asn Gln Thr Asn Lys Leu Val Asp Asp Gly Lys Lys Lys Pro Arg
850 855 860

Ala Lys Gln Lys Thr Asn Glu Lys Gly Thr Ser Arg Lys His Thr Thr
865 870 875 880

Leu Lys Asp Glu Lys Ile Lys Ser Gln Ser Gly Ala Glu Val Lys Phe
885 890 895

Val Leu Lys His Gln Asn Val Ser Glu Phe Ala Ser Ser Ser Gly Gly
900 905 910

Ser Gln Leu Leu Phe Lys Gln Lys Asp Met Pro Leu Met Gly Ser Ala
915 920 925

Val Asp His Pro Leu Ser Ala Ser Leu Pro Thr Gly Ile Asn Ala Gln
930 935 940

Gln Lys Leu Ser Gly Cys Phe Ser Ser Phe Leu Glu Ser Lys Lys Ser
945 950 955 960

Val Asp Leu Gln Thr Phe Pro Ser Ser Arg Asp Asp Leu His Pro Ser
965 970 975

Val Val Cys Asn Ser Ile Gly Pro Gly Val Ser Lys Ile Asn Val Gln
980 985 990

Arg Pro His Asn Gln Ser Ala Met Phe Thr Leu Lys Glu Ser Thr Leu
995 1000 1005

Ile Gln Lys Asn Ile Phe Asp Leu Ser Asn His Leu Ser Gln Val
1010 1015 1020

Ala Gln Asn Thr Gln Ile Ser Ser Gly Met Ser Ser Lys Ile Glu
1025 1030 1035

Asp Asn Ala Asn Asn Ile Gln Arg Asn Tyr Leu Ser Ser Ile Gly

1040 1045 1050

Lys Leu Ser Glu Tyr Arg Asn Ser Leu Glu Ser Lys Leu Asp Gln
1055 1060 1065

Ala Tyr Thr Pro Asn Phe Leu His Cys Lys Asp Ser Gln Gln Gln
1070 1075 1080

Ile Val Cys Ile Ala Glu Gln Ser Lys His Ser Glu Thr Cys Ser
1085 1090 1095

Pro Gly Asn Thr Ala Ser Glu Glu Ser Gln Met Pro Asn Asn Cys
1100 1105 1110

Phe Val Thr Ser Leu Arg Ser Pro Ile Lys Gln Ile Ala Trp Glu
1115 1120 1125

Gln Lys Gln Arg Gly Phe Ile Leu Asp Met Ser Asn Phe Lys Pro
1130 1135 1140

Glu Arg Val Lys Pro Arg Ser Leu Ser Glu Ala Ile Ser Gln Thr
1145 1150 1155

Lys Ala Leu Ser Gln Cys Lys Asn Arg Asn Val Ser Thr Pro Ser
1160 1165 1170

Ala Phe Gly Glu Gly Gln Ser Gly Leu Ala Val Leu Lys Glu Leu
1175 1180 1185

Leu Gln Lys Arg Gln Gln Lys Ala Gln Asn Ala Asn Thr Thr Gln
1190 1195 1200

Asp Pro Leu Ser Asn Lys His Gln Pro Asn Lys Asn Ile Ser Gly
1205 1210 1215

Ser Leu Glu His Asn Lys Ala Asn Lys Arg Thr Arg Ser Val Thr
1220 1225 1230

Ser Pro Arg Lys Pro Arg Thr Pro Arg Ser Thr Lys Gln Lys Glu
1235 1240 1245

Lys Ile Pro Lys Leu Leu Lys Val Asp Ser Leu Asn Leu Gln Asn
1250 1255 1260

Ser Ser Gln Leu Asp Asn Ser Val Ser Asp Asp Ser Pro Ile Phe
1265 1270 1275

Phe Ser Asp Pro Gly Phe Glu Ser Cys Tyr Ser Leu Glu Asp Ser
1280 1285 1290

Leu Ser Pro Glu His Asn Tyr Asn Phe Asp Ile Asn Thr Ile Gly
1295 1300 1305

Gln Thr Gly Phe Cys Ser Phe Tyr Ser Gly Ser Gln Phe Val Pro
1310 1315 1320

Ala Asp Gln Asn Leu Pro Gln Lys Phe Leu Ser Asp Ala Val Gln
1325 1330 1335

Asp Leu Phe Pro Gly Gln Ala Ile Glu Lys Asn Glu Phe Leu Ser
1340 1345 1350

His Asp Asn Gln Lys Cys Asp Glu Asp Lys His His Thr Thr Asp
1355 1360 1365

Ser Ala Ser Trp Ile Arg Ser Gly Thr Leu Ser Pro Glu Ile Phe
1370 1375 1380

Glu Lys Ser Thr Ile Asp Ser Asn Glu Asn Arg Arg His Asn Gln
1385 1390 1395

Trp Lys Asn Ser Phe His Pro Leu Thr Thr Arg Ser Asn Ser Ile
1400 1405 1410

Met Asp Ser Phe Cys Val Gln Gln Ala Glu Asp Cys Leu Ser Glu
1415 1420 1425

Lys Ser Arg Leu Asn Arg Ser Ser Val Ser Lys Glu Val Phe Leu
1430 1435 1440

Ser Leu Pro Gln Pro Asn Asn Ser Asp Trp Ile Gln Gly His Thr
1445 1450 1455

Arg Lys Glu Met Gly Gln Ser Leu Asp Ser Ala Asn Thr Ser Phe
1460 1465 1470

Thr Ala Ile Leu Ser Ser Pro Asp Gly Glu Leu Val Asp Val Ala
1475 1480 1485

Cys Glu Asp Leu Glu Leu Tyr Val Ser Arg Asn Asn Asp Met Leu
1490 1495 1500

Thr Pro Thr Pro Asp Ser Ser Pro Arg Ser Thr Ser Ser Pro Ser
1505 1510 1515

Gln Ser Lys Asn Gly Ser Phe Thr Pro Arg Thr Ala Asn Ile Leu
1520 1525 1530

Lys Pro Leu Met Ser Pro Pro Ser Arg Glu Glu Ile Met Ala Thr
1535 1540 1545

Leu Leu Asp His Asp Leu Ser Glu Thr Ile Tyr Gln Glu Pro Phe
1550 1555 1560

Cys Ser Asn Pro Ser Asp Val Pro Glu Lys Pro Arg Glu Ile Gly
1565 1570 1575

Gly Arg Leu Leu Met Val Glu Thr Arg Leu Ala Asn Asp Leu Ala
1580 1585 1590

Glu Phe Glu Gly Asp Phe Ser Leu Glu Gly Leu Arg Leu Trp Lys
1595 1600 1605

Thr Ala Phe Ser Ala Met Thr Gln Asn Pro Arg Pro Gly Ser Pro
1610 1615 1620

Leu Arg Ser Gly Gln Gly Val Val Asn Lys Gly Ser Ser Asn Ser
1625 1630 1635

Pro Lys Met Val Glu Asp Lys Lys Ile Val Ile Met Pro Cys Lys
1640 1645 1650

Cys Ala Pro Ser Arg Gln Leu Val Gln Val Trp Leu Gln Ala Lys
1655 1660 1665

Glu Glu Tyr Glu Arg Ser Lys Lys Leu Pro Lys Thr Lys Pro Thr
1670 1675 1680

Gly Val Val Lys Ser Ala Glu Asn Phe Ser Ser Ser Val Asn Pro
1685 1690 1695

Asp Asp Lys Pro Val Val Pro Pro Lys Met Asp Val Ser Pro Cys
1700 1705 1710

Ile Leu Pro Thr Thr Ala His Thr Lys Glu Asp Val Asp Asn Ser
1715 1720 1725

Gln Ile Ala Leu Gln Ala Pro Thr Thr Gly Cys Ser Gln Thr Ala
1730 1735 1740

Ser Glu Ser Gln Met Leu Pro Pro Val Ala Ser Ala Ser Asp Pro
1745 1750 1755

Glu Lys Asp Glu Asp Asp Asp Asp Asn Tyr Tyr Ile Ser Tyr Ser
1760 1765 1770

Ser Pro Asp Ser Pro Val Ile Pro Pro Trp Gln Gln Pro Ile Ser
1775 1780 1785

Pro Asp Ser Lys Ala Leu Asn Gly Asp Asp Arg Pro Ser Ser Pro
1790 1795 1800

Val Glu Glu Leu Pro Ser Leu Ala Phe Glu Asn Phe Leu Lys Pro
1805 1810 1815

Ile Lys Asp Gly Ile Gln Lys Ser Pro Cys Ser Glu Pro Gln Glu
1820 1825 1830

Pro Leu Val Ile Ser Pro Ile Asn Thr Arg Ala Arg Thr Gly Lys
1835 1840 1845

Cys Glu Ser Leu Cys Phe His Ser Thr Pro Ile Ile Gln Arg Lys
1850 1855 1860

Leu Leu Glu Arg Leu Pro Glu Ala Pro Gly Leu Ser Pro Leu Ser
1865 1870 1875

Thr Glu Pro Lys Thr Gln Lys Leu Ser Asn Lys Lys Gly Ser Asn
1880 1885 1890

Thr Asp Thr Leu Arg Arg Val Leu Leu Thr Gln Ala Lys Asn Gln
1895 1900 1905

Phe Ala Ala Val Asn Thr Pro Gln Lys Glu Thr Ser Gln Ile Asp
1910 1915 1920

Gly Pro Ser Leu Asn Asn Thr Tyr Gly Phe Lys Val Ser Ile Gln
1925 1930 1935

Asn Leu Gln Glu Ala Lys Ala Leu His Glu Ile Gln Asn Leu Thr
1940 1945 1950

Leu Ile Ser Val Glu Leu His Ala Arg Thr Arg Arg Asp Leu Glu

1955 1960 1965

Pro Asp Pro Glu Phe Asp Pro Ile Cys Ala Leu Phe Tyr Cys Ile
1970 1975 1980

Ser Ser Asp Thr Pro Leu Pro Asp Thr Glu Lys Thr Glu Leu Thr
1985 1990 1995

Gly Val Ile Val Ile Asp Lys Asp Lys Thr Val Phe Ser Gln Asp
2000 2005 2010

Ile Arg Tyr Gln Thr Pro Leu Leu Ile Arg Ser Gly Ile Thr Gly
2015 2020 2025

Leu Glu Val Thr Tyr Ala Ala Asp Glu Lys Ala Leu Phe His Glu
2030 2035 2040

Ile Ala Asn Ile Ile Lys Arg Tyr Asp Pro Asp Ile Leu Leu Gly
2045 2050 2055

Tyr Glu Ile Gln Met His Ser Trp Gly Tyr Leu Leu Gln Arg Ala
2060 2065 2070

Ala Ala Leu Ser Ile Asp Leu Cys Arg Met Ile Ser Arg Val Pro
2075 2080 2085

Asp Asp Lys Ile Glu Asn Arg Phe Ala Ala Glu Arg Asp Glu Tyr
2090 2095 2100

Gly Ser Tyr Thr Met Ser Glu Ile Asn Ile Val Gly Arg Ile Thr
2105 2110 2115

Leu Asn Leu Trp Arg Ile Met Arg Asn Glu Val Ala Leu Thr Asn
2120 2125 2130

Tyr Thr Phe Glu Asn Val Ser Phe His Val Leu His Gln Arg Phe
2135 2140 2145

Pro Leu Phe Thr Phe Arg Val Leu Ser Asp Trp Phe Asp Asn Lys
2150 2155 2160

Thr Asp Leu Tyr Arg Tyr Cys Ser Ile Thr Leu Lys Lys Arg Gln
2165 2170 2175

Gln Thr Ser Ala Leu Tyr His Trp Gln Val Leu Gly Pro Ile Tyr
2180 2185 2190

Phe Trp Val Ile Phe Thr Ser Tyr Asn Ile Lys Ile Leu Phe Met
2195 2200 2205

Asp Leu Leu Arg Val Leu Leu Phe Val Phe Leu Arg Arg Trp Lys
2210 2215 2220

Met Val Asp His Tyr Val Ser Arg Val Arg Gly Asn Leu Gln Met
2225 2230 2235

Leu Glu Gln Leu Asp Leu Ile Gly Lys Thr Ser Glu Met Ala Arg
2240 2245 2250

Leu Phe Gly Ile Gln Phe Leu His Val Leu Thr Arg Gly Ser Gln
2255 2260 2265

Tyr Arg Val Glu Ser Met Met Leu Arg Ile Ala Lys Pro Met Asn
2270 2275 2280

Tyr Ile Pro Val Thr Pro Ser Val Gln Gln Arg Ser Gln Met Arg
2285 2290 2295

Ala Pro Gln Cys Val Pro Leu Ile Met Glu Pro Glu Ser Arg Phe
2300 2305 2310

Tyr Ser Asn Ser Val Leu Val Leu Asp Phe Gln Ser Leu Tyr Pro
2315 2320 2325

Ser Ile Val Ile Ala Tyr Asn Tyr Cys Phe Ser Thr Cys Leu Gly
2330 2335 2340

His Val Glu Asn Leu Gly Lys Tyr Asp Glu Phe Lys Phe Gly Cys
2345 2350 2355

Thr Ser Leu Arg Val Pro Pro Asp Leu Leu Tyr Gln Val Arg His
2360 2365 2370

Asp Ile Thr Val Ser Pro Asn Gly Val Ala Phe Val Lys Pro Ser
2375 2380 2385

Val Arg Lys Gly Val Leu Pro Arg Met Leu Glu Glu Ile Leu Lys
2390 2395 2400

Thr Arg Phe Met Val Lys Gln Ser Met Lys Ala Tyr Lys Gln Asp
2405 2410 2415

Arg Ala Leu Ser Arg Met Leu Asp Ala Arg Gln Leu Gly Leu Lys
2420 2425 2430

Leu Ile Ala Asn Val Thr Phe Gly Tyr Thr Ser Ala Asn Phe Ser
2435 2440 2445

Gly Arg Met Pro Cys Ile Glu Val Gly Asp Ser Ile Val His Lys
2450 2455 2460

Ala Arg Glu Thr Leu Glu Arg Ala Ile Lys Leu Val Asn Asp Thr
2465 2470 2475

Lys Lys Trp Gly Ala Arg Val Val Tyr Gly Asp Thr Asp Ser Met
2480 2485 2490

Phe Val Leu Leu Lys Gly Ala Thr Lys Glu Gln Ser Phe Lys Ile
2495 2500 2505

Gly Gln Glu Ile Ala Glu Ala Val Thr Ala Thr Asn Pro Lys Pro
2510 2515 2520

Val Lys Leu Lys Phe Glu Lys Val Tyr Leu Pro Cys Val Leu Gln
2525 2530 2535

Thr Lys Lys Arg Tyr Val Gly Tyr Met Tyr Glu Thr Leu Asp Gln
2540 2545 2550

Lys Asp Pro Val Phe Asp Ala Lys Gly Ile Glu Thr Val Arg Arg
2555 2560 2565

Asp Ser Cys Pro Ala Val Ser Lys Ile Leu Glu Arg Ser Leu Lys
2570 2575 2580

Leu Leu Phe Glu Thr Arg Asp Ile Ser Leu Ile Lys Gln Tyr Val
2585 2590 2595

Gln Arg Gln Cys Met Lys Leu Leu Glu Gly Lys Ala Ser Ile Gln
2600 2605 2610

Asp Phe Ile Phe Ala Lys Glu Tyr Arg Gly Ser Phe Ser Tyr Lys
2615 2620 2625

Pro Gly Ala Cys Val Pro Ala Leu Glu Leu Thr Ser Phe Phe Ile
2630 2635 2640

Val Leu Leu Leu Phe Asn Ser Asp Leu Ile Cys Glu Lys Asp Gly
2645 2650 2655

Phe His Asn Ser Ile Trp Val Trp Phe Phe Ser Leu Asn Ser Asn
2660 2665 2670

Arg Lys Met Leu Thr Tyr Asp Arg Arg Ser Glu Pro Gln Val Gly
2675 2680 2685

Glu Arg Val Pro Tyr Val Ile Ile Tyr Gly Thr Pro Gly Val Pro
2690 2695 2700

Leu Ile Gln Leu Val Arg Arg Pro Val Glu Val Leu Gln Asp Pro
2705 2710 2715

Thr Leu Arg Leu Asn Ala Thr Tyr Tyr Ile Thr Lys Gln Ile Leu
2720 2725 2730

Pro Pro Leu Ala Arg Ile Phe Ser Leu Ile Gly Ile Asp Val Phe
2735 2740 2745

Ser Trp Tyr His Glu Leu Pro Arg Ile His Lys Ala Thr Ser Ser
2750 2755 2760

Ser Arg Ser Glu Pro Glu Gly Arg Lys Gly Thr Ile Ser Gln Tyr
2765 2770 2775

Phe Thr Thr Leu His Cys Pro Val Cys Asp Asp Leu Thr Gln His
2780 2785 2790

Gly Ile Cys Ser Lys Cys Arg Ser Gln Pro Gln His Val Ala Val
2795 2800 2805

Ile Leu Asn Gln Glu Ile Arg Glu Leu Glu Arg Gln Gln Glu Gln
2810 2815 2820

Leu Val Lys Ile Cys Lys Asn Cys Thr Gly Cys Phe Asp Arg His
2825 2830 2835

Ile Pro Cys Val Ser Leu Asn Cys Pro Val Leu Phe Lys Leu Ser
2840 2845 2850

Arg Val Asn Arg Glu Leu Ser Lys Ala Pro Tyr Leu Arg Gln Leu
2855 2860 2865

Leu Asp Gln Phe

2870

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
2 September 2004 (02.09.2004)

PCT

(10) International Publication Number
WO 2004/073657 A3

(51) International Patent Classification[7]: C12Q 1/68, C12N 15/52

(21) International Application Number: PCT/US2004/005455

(22) International Filing Date: 19 February 2004 (19.02.2004)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/448,784 19 February 2003 (19.02.2003) US

(71) Applicant *(for all designated States except US)*: **PROTEIN DESIGN LABS, INC.** [US/US]; 34801 Campus Drive, Fremont, CA 94555 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **AZIZ, Natasha** [US/US]; 411 California Avenue #5, Palo Alto, CA 94306 (US). **GISH, Kurt, C.** [US/US]; 37 Artuna Avenue, Piedmont, CA 94611 (US). **WILSON, Keith, E.** [TZ/US]; 219 Jeter Street, Redwood City, CA 94062 (US). **ZLOTNIK, Albert** [US/US]; 507 Alger Drive, Palo Alto, CA 94306 (US).

(74) Agents: **HALLUIN, Albert** et al.; 301 Ravenswood Avenue, Menlo Park, CA 94025 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

(88) Date of publication of the international search report:
21 April 2005

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: METHODS OF DIAGNOSIS OF CANCER AND OTHER DISEASES, COMPOSITION AND METHODS OF SCREENING FOR MODULATORS OF CANCER AND OTHER DISEASES

(57) Abstract: Described herein are genes whose expression are up-regulated or down-regulated in specific cancers or other diseases, or are otherwise regulated in disease. Related methods and compositions that can be used for diagnosis, prognosis, and treatment of those medical conditions are disclosed. Also described herein are methods that can be used to identify modulators of these selected conditions.

WO 2004/073657 A3

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US2004/005455

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12Q1/68 C12N15/52

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C12Q C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
EPO-Internal, Sequence Search, WPI Data, PAJ, BIOSIS, EMBASE

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2002/086389 A1 (CORLEY NEIL C ET AL) 4 July 2002 (2002-07-04) claims 27,28,30,48-55	1-27
A	US 6 479 241 B1 (ALLER ALEX) 12 November 2002 (2002-11-12) column 3, lines 4-6; claim 1; table 1	1-27
A	WO 01/32693 A (PELLETIER JERRY ; PRAWITT DIRK (DE); ZABEL BERNHARD (DE); JOHANNES GUT) 10 May 2001 (2001-05-10) claims 13,15; examples 6,7	1-27
A	WO 97/39139 A (SMITHKLINE BEECHAM CORP ; ROBBINS DAVID J (US)) 23 October 1997 (1997-10-23) claim 1	1-27

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
26 October 2004	11 01 2005

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL – 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Knudsen, H

Form PCT/ISA/210 (second sheet) (January 2004)

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US2004/005455

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	EP 1 136 547 A (PFIZER PROD INC) 26 September 2001 (2001-09-26) paragraphs '0072!, '0076!, '0077!, '0087!; claims 1,11,14,15; example 1	1-27
X	DATABASE GENESEQ EBI; Human breast cancer expressed polynucleotide 19 July 2001 (2001-07-19), XP002296526 Database accession no. AAL26571 abstract	1
X	WO 02/42439 A (INST GENETICS LLC) 30 May 2002 (2002-05-30) page 7, lines 3-6; sequences 2,4-6,8 page 17, paragraph 2	17,22,23
X	WO 00/53744 A (DIVERSA CORP) 14 September 2000 (2000-09-14) claims 2,20,21,32-35; figure 9	17,23, 26,27
A	WO 01/83782 A (PLOWMAN GREGORY D ; PAYNE VILIA (US); SUGEN INC (US); WHYTE DAVID (US)) 8 November 2001 (2001-11-08) page 127, paragraph 3; claims 9,12-14,24,25 page 138, paragraph 4 page 155, paragraph 2 page 179, paragraph 2	1-27
P,A	LLAMAZARES MARIA ET AL: "Identification and characterization of ADAMTS-20 defines a novel subfamily of metalloproteinases-disintegrins with multiple thrombospondin-1 repeats and a unique GON domain." THE JOURNAL OF BIOLOGICAL CHEMISTRY. 11 APR 2003, vol. 278, no. 15, 11 April 2003 (2003-04-11), pages 13382-13389, XP002296532 ISSN: 0021-9258 abstract	1-27

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2004/005455

Box II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

Although claims 22 and 24 are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

2. [ ] Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. [ ] As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. [ ] As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. [X] No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

1-27(partially)

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

Invention 1: Claims 1-27 (partially)

Method for detecting a pathological cell in a biological sample by detection of a nucleic acid with SEQ ID NO 1 or a protein with SEQ ID NO 59. An isolated nucleic acid with SEQ ID NO 1. Vectors or host cells containing the isolated nucleic acid with SEQ ID NO:1. An isolated nucleic acid encoding the polypeptide with SEQ ID NO:59. A polypeptide encoded by the nucleic with SEQ ID NO:1. An antibody that binds specifically the polypeptide with SEQ ID NO:1. A method for targetting a pathological cell in a patient employing the said antibody. A method for detecting a pathological cell in a patient employing said antibody. A method for identifying compounds modulating pathology-associated polypeptides by contacting the compound with a polypeptide encoded by the polynucleotide with SEQ ID NO:1. Screening assay involving the comparison of expression levels of SEQ ID NO:1 from a pathological cell and a cell that does not show the pathology.

Invention 2: Claims 1-27 (partially)

Method for detecting a pathological cell in a biological sample by detection of a nucleic acid with SEQ ID NO 2 or a protein with SEQ ID NO 60. An isolated nucleic acid with SEQ ID NO 2. Vectors or host cells containing the isolated nucleic acid with SEQ ID NO:2. An isolated nucleic acid encoding the polypeptide with SEQ ID NO:60. A polypeptide encoded by the nucleic with SEQ ID NO:2. An antibody that binds specifically the polypeptide with SEQ ID NO:2. A method for targetting a pathological cell in a patient employing the said antibody. A method for detecting a pathological cell in a patient employing said antibody. A method for identifying compounds modulating pathology-associated polypeptides by contacting the compound with a polypeptide encoded by the polynucleotide with SEQ ID NO:2. Screening assay involving the comparison of expression levels of SEQ ID NO:2 from a pathological cell and a cell that does not show the pathology.

Inventions 3-58

idem for each of the polynucleotide - polypeptide pairs with SEQ ID NOs 3-58 and 61-116, as defined in Table 2.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/US2004/005455

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2002086389 A1	04-07-2002	US 6379722 B1	30-04-2002
		US 5876996 A	02-03-1999
US 6479241 B1	12-11-2002	NONE	
WO 0132693 A	10-05-2001	DE 19953167 A1	26-07-2001
		AU 2347501 A	14-05-2001
		WO 0132693 A2	10-05-2001
		EP 1237910 A2	11-09-2002
WO 9739139 A	23-10-1997	AU 2676197 A	07-11-1997
		EP 0904397 A1	31-03-1999
		JP 2000510331 T	15-08-2000
		WO 9739139 A1	23-10-1997
		ZA 9703231 A	25-11-1997
EP 1136547 A	26-09-2001	EP 1136547 A2	26-09-2001
		JP 2001327297 A	27-11-2001
		US 2002090373 A1	11-07-2002
WO 0242439 A	30-05-2002	AU 3968002 A	03-06-2002
		WO 0242439 A2	30-05-2002
WO 0053744 A	14-09-2000	US 6238884 B1	29-05-2001
		US 6352842 B1	05-03-2002
		US 6537776 B1	25-03-2003
		AU 3879300 A	28-09-2000
		CA 2361927 A1	14-09-2000
		EP 1161529 A2	12-12-2001
		JP 2002537836 A	12-11-2002
		WO 0053744 A2	14-09-2000
		US 2004152077 A1	05-08-2004
		US 2002142394 A1	03-10-2002
		US 2003194763 A1	16-10-2003
		US 2003036116 A1	20-02-2003
		US 2003219752 A1	27-11-2003
		US 2003207287 A1	06-11-2003
		US 2004029174 A1	12-02-2004
		US 6713279 B1	30-03-2004
		US 6358709 B1	19-03-2002
		US 6361974 B1	26-03-2002
		US 2004248143 A1	09-12-2004
		US 2002119457 A1	29-08-2002
		US 2002146762 A1	10-10-2002
		AU 4039400 A	16-10-2000
		CA 2329122 A1	05-10-2000
		EP 1092041 A1	18-04-2001
		JP 2004500019 T	08-01-2004
		WO 0058517 A1	05-10-2000
		US 6479258 B1	12-11-2002
		AU 5624600 A	02-01-2001
		CA 2374667 A1	21-12-2000
		EP 1192280 A1	03-04-2002
		JP 2004500027 T	08-01-2004
		WO 0077262 A1	21-12-2000
		US 2004002103 A1	01-01-2004
		US 6605449 B1	12-08-2003

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/US2004/005455

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 0183782 A	08-11-2001	AU 5947301 A	12-11-2001
		CA 2408105 A1	08-11-2001
		EP 1294901 A2	26-03-2003
		JP 2004504812 T	19-02-2004
		WO 0183782 A2	08-11-2001